

UNIVERSIDAD COMPLUTENSE DE MADRID

**FACULTAD DE MEDICINA
DEPARTAMENTO DE MICROBIOLOGÍA I**



TESIS DOCTORAL

**Estudios multicéntricos de la esclerosis múltiple:
buscando la heredabilidad perdida**

**PRESENTADA POR
Miguel Ángel Ortiz Rosales**

Directoras

Elena Urcelay García
Concepción Núñez Pardo de Vera

Madrid, 2015

La *Dra. Elena Urcelay García*, investigadora estabilizada de la Comunidad de Madrid,
y la *Dra. M^a Concepción Núñez Pardo de Vera*, investigadora del Sistema Nacional de
Salud “Miguel Servet II”,

CERTIFICAN:

Que el trabajo titulado **“Estudios genéticos multicéntricos de la esclerosis múltiple: buscando la heredabilidad perdida”** ha sido realizado bajo su dirección en el Servicio de Inmunología Clínica del Hospital Clínico San Carlos de Madrid por D. Miguel Ángel Ortiz Rosales, licenciado en Biología por la Universidad Complutense de Madrid, para optar al GRADO DE DOCTOR por la misma universidad.

Madrid, 8 de mayo de 2015.

Elena Urcelay García

M^a Concepción Núñez Pardo de Vera



Informe del Director de la Tesis Doctoral

DATOS DE LA TESIS DOCTORAL	
Nombre del Doctorando	MIGUEL ÁNGEL ORTIZ ROSALES
Título de la Tesis	ESTUDIOS GENÉTICOS MULTICÉNTRICOS DE LA ESCLEROSIS MÚLTIPLE: BUSCANDO LA HEREDABILIDAD PERDIDA
Facultad o Centro	FACULTAD DE MEDICINA (UCM)

DATOS DEL DIRECTOR DE LA TESIS DOCTORAL	
Nombre Completo	M ^a Elena Urcelay García
Centro al que pertenece en la actualidad	Instituto de Investigación Sanitaria del H. Clínico San Carlos-IdISSC
D.N.I./Pasaporte	07219750
e-mail	elena.urcelay@salud.madrid.org

	VALORACIÓN DE LA TESIS			
	Muy Buena	Buena	Suficiente	Deficiente
Originalidad		x		
Definición Objetivos	x			
Metodología	x			
Relevancia Resultados	x			
Discusión / Conclusiones	x			

INFORME (en caso necesario se podrán añadir más hojas):

Los objetivos de esta tesis son realistas y está planteada y desarrollada con rigor, en ocasiones en contexto de colaboraciones con grupos nacionales e internacionales. La metodología es adecuada al volumen de enfermos estudiados. Los resultados obtenidos contribuirán sin duda a mejorar el conocimiento de la esclerosis múltiple y, como consecuencia, a incrementar la calidad de vida de los afectados. Las conclusiones dan respuesta al planteamiento inicial y la bibliografía está bien actualizada.

Es, por tanto, un trabajo de factura impecable con una presentación clara a pesar de la dificultad inherente a los datos.

Madrid, a 9 de Abril de 2015

Fdo.:

Este impreso deberá entregarse a la Comisión Académica responsable del Programa de Doctorado, para su estudio y aprobación en la admisión a trámite de la tesis doctoral, con el fin de que se incluya junto con la documentación enviada a la Comisión de Doctorado para la aprobación, si procede, y designación del Tribunal que ha de juzgar la Tesis Doctoral en el acto de defensa.



Informe del Director de la Tesis Doctoral

DATOS DE LA TESIS DOCTORAL	
Nombre del Doctorando	MIGUEL ÁNGEL ORTIZ ROSALES
Título de la Tesis	ESTUDIOS GENÉTICOS MULTICÉNTRICOS DE LA ESCLEROSIS MÚLTIPLE: BUSCANDO LA HEREDABILIDAD PERDIDA
Facultad o Centro	FACULTAD DE MEDICINA (UCM)

DATOS DEL DIRECTOR DE LA TESIS DOCTORAL	
Nombre Completo	M Concepción Núñez Pardo de Vera
Centro al que pertenece en la actualidad	Instituto de Investigación Sanitaria del Hospital Clínico San Carlos
D.N.I./Pasaporte	33299125R
e-mail	mariaconcepcion.nunez@salud.madrid.org

	VALORACIÓN DE LA TESIS			
	Muy Buena	Buena	Suficiente	Deficiente
Originalidad		x		
Definición Objetivos		x		
Metodología	X			
Relevancia Resultados	X			
Discusión / Conclusiones		x		

INFORME (en caso necesario se podrán añadir más hojas):

La tesis presentada por M Ángel Ortiz presenta unos objetivos bien enunciados que son tratados de manera clara y con la metodología adecuada en el desarrollo de sus trabajos. Los resultados obtenidos se interpretan de manera correcta, siendo cautos, cuando es necesario, con respecto a las conclusiones que se pueden extraer.

En general, creo que el doctorando demuestra tener una buena visión general de la problemática que aborda en su tesis doctoral y que sus trabajos permiten avanzar en el conocimiento de las bases genéticas de la esclerosis múltiple.

Madrid, a 30 de Abril de 2015

Fdo.:

Este impreso deberá entregarse a la Comisión Académica responsable del Programa de Doctorado, para su estudio y aprobación en la admisión a trámite de la tesis doctoral, con el fin de que se incluya junto con la documentación enviada a la Comisión de Doctorado para la aprobación, si procede, y designación del Tribunal que ha de juzgar la Tesis Doctoral en el acto de defensa.

La realización de este trabajo se ha visto facilitada gracias a la financiación otorgada por el Fondo de Investigaciones Sanitarias (FEDER-FIS) del Instituto de Salud Carlos III (PI10/01985) y por la Fundación Alicia Koplowitz (Descifrando las claves de la esclerosis múltiple: de la genética a la etiología).

Dedicada a mis padres.

Índice



ABREVIATURAS.....	5
RESUMEN/SUMMARY	11
INTRODUCTION.....	21
OBJECTIVES	21
MATERIALS AND METHODS	22
RESULTS AND DISCUSSION	25
CONCLUSIONS	27
INTRODUCCIÓN	29
1. DEFINICIÓN Y EPIDEMIOLOGÍA	31
2. FORMAS CLÍNICAS	32
3. DIAGNÓSTICO.....	33
4. INMUNOPATOGENÉISIS DE LA ESCLEROSIS MÚLTIPLE	35
4.1. <i>Procesos inmunológicos en la esclerosis múltiple</i>	36
4.2. <i>Componentes celulares de la respuesta inmunitaria en la esclerosis múltiple</i>	40
5. TRATAMIENTO.....	43
5.1. <i>Propiedades inmunomoduladoras</i>	44
5.2. <i>Inhibición de la migración linfocitaria</i>	44
5.3. <i>Propiedades inmunosupresoras</i>	45
5.4. <i>Efecto sobre los linfocitos T</i>	45
5.5. <i>Efecto sobre los linfocitos B</i>	45
6. ETIOLOGÍA DE LA ESCLEROSIS MÚLTIPLE	45
6.1. <i>Factores de riesgo ambientales</i>	46
6.2. <i>Factores de riesgo genéticos</i>	48
7. BÚSQUEDA DE LA “HEREDABILIDAD PERDIDA” EN LA ESCLEROSIS MÚLTIPLE	61
OBJETIVOS.....	67
MATERIALES Y MÉTODOS.....	71
1. PACIENTES Y CONTROLES	73
2. PREPARACIÓN DE LAS MUESTRAS: AISLAMIENTO, CUANTIFICACIÓN Y ESTIMACIÓN DE LA PUREZA DEL DNA	74
3. GENOTIPADO DE LAS MUESTRAS.....	74
4. DISEÑO DE LOS ESTUDIOS MULTICÉNTRICOS	76
4.1. <i>Pacientes de esclerosis múltiple y controles</i>	76
4.2. <i>Búsqueda y abstracción de datos</i>	78
4.3. <i>Análisis de los datos procedentes de los GWAS</i>	79
4.4. <i>Selección de SNPs</i>	80
4.5. <i>Genotipado de SNPs</i>	84
4.6. <i>Controles de calidad</i>	86
5. ESTRATEGIA DE GENOTIPADO DEL GEN LILRA3	87
6. MUESTRAS PARA SEPARACIÓN DE CÉLULAS ACTIVADAS POR FLUORESCENCIA (FACS) Y ELISA	87

7.	CITOMETRÍA DE FLUJO PARA ESTUDIO DE EXPRESIÓN EN SUPERFICIE CELULAR	88
8.	ENSAYOS DE CULTIVO CELULAR, PROLIFERACIÓN Y ELISA PARA LA CUANTIFICACIÓN DE CITOQUINAS	90
9.	ANÁLISIS DE DATOS	90
9.1.	<i>Potencia y corrección múltiple</i>	90
9.2.	<i>Estudios caso-control</i>	91
9.3.	<i>Estudio de haplotipos</i>	92
9.4.	<i>Estudio de interacciones</i>	92
9.5.	<i>Análisis de muestras procedentes de distintas poblaciones</i>	93
9.6.	<i>Análisis estadístico de los datos del FACS</i>	95
RESULTADOS Y DISCUSIÓN.....		97
1.	ESTUDIO DE CRIBADO DE GENES DE CANDIDATOS Y VALIDACIÓN DE LAS SEÑALES SIGNIFICATIVAS.....	99
2.	ESTUDIO DE ASOCIACIÓN A GRAN ESCALA DEL GEN APOE EN ESCLEROSIS MÚLTIPLE	109
3.	FOLLOW-UP DEL GWAS DESARROLLADO POR EL IMSGC Y ESTUDIO DE OTROS GENES CANDIDATOS	115
4.	ESTUDIO DE CINCO GENES SUGERENTES DE ASOCIACIÓN TRAS GWAS.....	123
5.	MAPEO FINO Y ANÁLISIS FUNCIONAL DEL GEN DE RIESGO CD6.....	129
5.1.	<i>Estudio de variantes de riesgo</i>	129
5.2.	<i>Determinación de CD6 en función del haplotipo rs11230563-rs2074225</i>	135
5.3.	<i>Estudio de la proliferación y secreción de citoquinas tras estimulación de linfocitos T los diferentes haplotipos</i>	139
6.	INFLUENCIA DE LA DELECCIÓN EN LILRA3 EN EL RIESGO A DESARROLLAR EM	142
ESTADO ACTUAL Y POSIBLES APROXIMACIONES EN EL ESTUDIO DE LOS FACTORES GENÉTICOS RELACIONADOS CON LA EM		147
CONCLUSIONES.....		153
BIBLIOGRAFÍA.....		157
ANEXO I: PUBLICACIONES		181

Abreviaturas

ADN	Ácido desoxirribonucleico
ALCAM	Molécula de adhesión celular de leucocitos activados (<i>Activated Leukocyte Cell Adhesion Molecule</i>)
ARN	Ácido ribonucleico
APOE	Apolipoproteína E
AR	Artritis reumatoide
BHE	Barrera hematoencefálica
CBLB	Proto-oncogen cbl-b (<i>Cbl proto-oncogen B</i>)
CDs	Células dendríticas (<i>Dendritic Cells</i>)
CD6	Cluster de diferenciación 6
CXCR5	Receptor 5 de quimiocinas con motivo C-X-C (<i>Chemokine C-X-C motif Receptor 5</i>)
DMT1	Diabetes mellitus tipo 1
EAE	Encefalitis autoinmune experimental
EC	Enfermedad celiaca
ELISA	Ensayo de inmunoabsorción ligado a enzimas (<i>Enzyme-Linked Immunoabsorbent Assay</i>)
EM	Esclerosis múltiple
EMRR	Esclerosis múltiple remitente recurrente
EMPP	Esclerosis múltiple primaria progresiva
EMPR	Esclerosis múltiple progresiva recurrente
EMSP	Esclerosis múltiple secundaria progresiva
FACS	Separación de células activadas por fluorescencia (<i>Fluorescence-Activated Cell Sorting</i>)
GWAS	Estudios de asociación del genoma completo (<i>Genome-Wide Association Studies</i>)
HLA	Antígeno leucocitario humano (<i>Human Leukocyte Antigen</i>)
HWE	Equilibrio de Hardy-Weinberg (<i>Hardy-Weinberg Equilibrium</i>)

IC	Intervalo de confianza
IFN	Interferón
IL12B	Interleuquina 12B
IMSGC	Consorcio Internacional sobre Genética de Esclerosis Múltiple (<i>International Multiple Sclerosis Genetics Consortium</i>)
kb	Kilobases
LB	Linfocito B
LCR	Líquido cefalorraquídeo
LD	Desequilibrio de ligamiento (<i>Linkage Disequilibrium</i>)
LES	Lupus eritematoso sistémico (<i>Systemic Lupus Erythematosus</i>)
LT	Linfocito T
LTh	Linfocito T cooperador
LT_{REG}	Linfocito T regulador
MAF	Frecuencia del alelo minoritario (<i>Minor Allele Frequency</i>)
MANBA	β -Manosidasa A lisosomal (<i>Mannosidase, β-A, lysosomal</i>)
MHC	Complejo principal de histocompatibilidad (<i>Major Histocompatibility Complex</i>)
mRNA	Ácido ribonucleico mensajero (<i>Messenger Ribonucleic Acid</i>)
NK	Células <i>natural killer</i>
OR	<i>Odds ratio</i>
pb	Pares de bases
PBMC	Células mononucleares de sangre periférica (<i>Peripheral Blood Mononuclear Cells</i>)
PCR	Reacción en cadena de la polimerasa (<i>Polymerase Chain Reaction</i>)
PVE	Potenciales visuales evocados
RM	Resonancia magnética
RPS6KB1	Proteína quinasa S6 ribosomal β -1 (<i>Ribosomal Protein S6 Kinase, 70 kDa, polypeptide 1</i>)

RT-PCR	Reacción en cadena de la polimerasa a tiempo real (<i>Real Time Polymerase Chain Reaction</i>)
SNC	Sistema nervioso central
SNP	Polimorfismo de un solo nucleótido (<i>Single Nucleotide Polymorphism</i>)
SOCS	Supresor de la señalización de citoquinas (<i>Suppresor of Cytokine Signaling</i>)
SOX8	Caja Y de la región determinante del sexo 8 (<i>Sex determining region Y-box 8</i>)
TME	Terapia modificadora de la enfermedad
TMEM39A	Proteína transmembrana 39A (<i>Transmembrane protein 39A</i>)
WTCCC	Consorcio de Casos y Controles del Wellcome Trust (<i>Wellcome Trust Case Control Consortium</i>)
VCAM-1	Molécula de adhesión vascular 1 (<i>Vascular Cell Adhesion Molecule 1</i>)
ZBTB46	Proteína 45 con dedos de zinc y un dominio BTB (<i>Zinc finger and BTB domain containing 46</i>)

Resumen/Summary

Introducción

La esclerosis múltiple (EM) es una enfermedad del sistema nervioso central caracterizada por desmielinización, inflamación multifocal, gliosis reactiva y pérdida oligodendrocitaria y axonal. La etiología de la EM es desconocida, sin embargo, se considera que es un trastorno multifactorial que se desencadena en individuos genéticamente susceptibles como consecuencia de la interacción con uno o varios factores ambientales.

El alelo *HLA-DRB1*15:01* es hasta el momento el principal factor genético asociado a EM pero existen diversos genes fuera de esta región que también parecen contribuir al riesgo de desarrollo de esta enfermedad, aunque con un efecto menor. Para identificar genes asociados fuera de la región HLA se han realizado estudios de ligamiento y estudios de asociación, estos últimos a través de estudios de genes candidatos y más recientemente con estudios de asociación mediante barrido genómico (GWAS). Sin embargo, existe un porcentaje de heredabilidad aún por descubrir (“heredabilidad perdida”) en el cual pueden estar involucrados distintos factores, algunos de los cuales abordamos en este trabajo.

Objetivos

1. Buscar genes de susceptibilidad a EM realizando un cribado de 368 polimorfismos localizados en genes de citoquinas, receptores de citoquinas, factores de transducción de la señal y receptores ionotrópicos y transportadores de glutamato; y validar las señales significativas en una cohorte independiente de replicación.
2. Realizar un estudio de asociación a gran escala de dos polimorfismos en el gen candidato *APOE* con datos procedentes de distintas fuentes.
3. Analizar la posible influencia en el riesgo a padecer EM de genes que se habían asociado en estudios previos.
4. Validar genes descritos como factores de riesgo a EM en estudios de GWAS y señales que quedaron al borde de la significación.

5. Identificar la variante o haplotipo causal de susceptibilidad a EM dentro del gen *CD6* y estudiar su papel funcional.
6. Buscar factores genéticos de susceptibilidad a EM no incluidos en los GWAS y que puedan contribuir a explicar la “heredabilidad perdida”: gen *LILRA3*.

Materiales y métodos

1. Pacientes y controles

Este trabajo incluye información de un máximo de 849 enfermos de EM y 832 controles de origen caucásico reclutados en el Hospital Clínico San Carlos.

2. Preparación de las muestras y método de genotipado

El aislamiento de DNA se realizó por el método de *salting out* a partir de los leucocitos procedentes de 10 ml de sangre periférica y el análisis de polimorfismos de un solo nucleótido (SNPs) se realizó mediante PCR a tiempo real utilizando sondas Taqman y empleando el equipo ABI PRISM 7900HT.

3. Diseño de los estudios multicéntricos

El presente trabajo se ha llevado a cabo en colaboración con distintos centros de investigación nacionales e internacionales. Para ello, se genotiparon muestras de hasta 9.457 casos y 10.691 controles descendientes directos de europeos procedentes de distintas poblaciones. En algún caso, se llevaron a cabo revisiones sistemáticas mediante búsquedas en la base de datos PubMed del NCBI y posterior abstracción de datos. También se solicitaron y obtuvieron datos de genotipos procedentes de las bases de datos de los GWAS, que fueron depurados usando *PLINK v.1.07* y posteriormente empleados para imputar nuevos datos mediante *IMPUTE v.2.0*.

4. Selección de SNPs

Para el estudio que incluye una fase inicial de cribado, se seleccionaron 368 SNPs en 55 genes o grupos de genes codificadores de citoquinas, receptores de citoquinas, supresores de la señalización de citoquinas, factores del complemento y receptores del glutamato. Para el

estudio de mapeo fino del gen *CD6*, se seleccionaron 13 SNPs marcadores de haplotipo y además se incluyó el SNP rs650258, localizado cerca del gen *CD6* y previamente asociado con la susceptibilidad a EM. En ambos casos, para la selección de SNPs se empleó el algoritmo de múltiples marcadores implementado en HapMap en una cohorte CEU (HapMap Release # 19).

En otros casos, la selección de SNPs se realizó en base a la información procedente de los GWAS en los que habían destacado como posibles factores de susceptibilidad a EM, por su asociación con otras enfermedades autoinmunes o por estudios de genes candidatos realizados previamente.

5. Genotipado de SNPs

Para estudiar los 368 SNPs localizados en genes de interés se diseñó un panel personalizado de SNPs en el Servicio de Atención al Cliente de *Illumina*. En los otros casos, el genotipado se realizó en el propio centro participante o en centros especializados de genotipado, tales como el Centro Nacional de Genotipado y el Instituto de Genética Molecular Max Planck, en función de las necesidades particulares del estudio o de la decisión de los investigadores de cada localización.

Debido al carácter multicéntrico de estos estudios, se realizaron diversas pruebas de control de calidad para garantizar que el genotipado de las muestras fuera concordante tanto dentro de un mismo centro como entre distintos centros participantes.

6. Estrategia de genotipado del gen *LILRA3*

Para estudiar la asociación entre la delección en el gen *LILRA3* y EM, se siguió un protocolo previamente descrito en el que se amplifican el alelo *wild type* y el alelo delecionado del gen *LILRA3* en un único tubo de PCR, los cuales se discriminan posteriormente por sus distintas movilidades electroforéticas en geles de agarosa.

7. Muestras para separación de células activadas por fluorescencia (FACS) y ELISA

Para este estudio se utilizaron 10 ml de sangre heparinizada en litio de 27 pacientes con EM y de 12 controles sanos. Las células mononucleares de sangre periférica (PBMC) se purificaron, se contaron e inmediatamente se utilizaron o se criopreservaron en la presencia de DMSO 10% (v/v) y 20% de suero fetal bovino, para posteriormente almacenarse a

-196°C hasta su utilización. Cuando se utilizaron PBMCs congeladas, éstas se descongelaron inmediatamente antes de su uso.

8. Citometría de flujo para estudio de expresión en superficie celular

Anticuerpos monoclonales anti-CD3 y anti-CD56 se usaron para identificar subgrupos de linfocitos, anti-CD45RA, anti-CD27 y anti-CD28 se usaron para distinguir células vírgenes, células de memoria centrales, células de memoria efectoras y células de memoria efectoras diferenciadas terminalmente. Para el marcaje, las PBMCs se descongelaron, se lavaron 2 veces y se incubaron durante 45 minutos. Posteriormente, las células se lavaron, se resuspendieron en 300 µl de tampón FACS y se analizaron en el citómetro FACS Canto. El software *FACS Diva v.6.1.3* se utilizó para analizar los datos crudos.

9. Ensayos de cultivo celular, proliferación y ELISA para la cuantificación de citoquinas

Para los ensayos de cultivo celular, proliferación y ELISA para la cuantificación de citoquinas, se resuspendieron 10⁶ PBMCs de cada donante en 500 µl de medio RPMI con 5% de suero fetal bovino, se alicuotaron en una placa de 48 pocillos y se cultivaron bajo 3 condiciones diferentes: sin estímulo y estimulado con anti-CD3 con o sin anticuerpo anti-CD6. Posteriormente, las PBMCs de los controles sanos se marcaron con 2 µM eFluor670 para determinar la proliferación celular. Después, se añadió medio estándar para detener el marcaje y se incubó en hielo durante 5 minutos. Las células marcadas se transfirieron a placas de 48 pocillos con el respectivo estímulo para los cultivos celulares.

Para estudiar el porcentaje de células que habían sufrido división a día 0 y día 3 por citometría de flujo, estas se incubaron con marcadores específicos de linaje. Los sobrenadantes de los cultivos se recogieron al tercer día y se cuantificaron los niveles de citoquinas IL-17 e IFN-γ por ELISA usando los kits *Ready-Set-Go*. La concentración de citoquinas (pg/ml) se calculó usando las curvas patrón generadas usando los respectivos calibradores.

10. Análisis de datos

La potencia se estimó usando el *Genetic Power Calculator* (<http://pngu.mgh.harvard.edu/~purcell/gpc/cc2.html>) o la calculadora de potencia de CATS (<http://www.sph.umich.edu/csg/abecasis/CaTS>). En el estudio caso-control se comparan las

frecuencias genotípicas, alélicas y de portadores aplicando el test χ^2 o el test exacto de Fisher (cuando los valores esperados fueron menores de 5), utilizando el programa *Statcalc* (*EpiInfo* v.6). Los análisis de haplotipos se realizaron en el programa *Haploview* o en el software *PLINK*. Los estudios de interacciones entre marcadores de genes no ligados y los análisis por género se realizaron mediante recuento directo a partir de tablas estratificadas. Para los meta-análisis se utilizaron los programas *Review Manager 5.0*, los paquetes estadísticos *HardyWeinberg v.1.4* y *rmeta v.2.16* en lenguaje R, el programa *PLINK* y el software *METAL*. Los análisis de regresión logística se realizaron en *PLINK*. Para calcular las diferencias en la expresión de moléculas de superficie se usaron pruebas t de *Student* y ANOVA de dos colas, según se comparasen dos o más factores, respectivamente. Las diferencias en la proliferación y la producción de citoquinas entre los diferentes haplotipos se estudió por la prueba no paramétrica U de Mann-Whitney en *Graphpad Prism v.5*.

Resultados y discusión

1. Estudio de cribado de genes candidatos y validación de las señales significativas

Las señales más significativas en la fase de cribado inicial se encontraron en los loci *SOCS1* ($P = 0,0005$), *IL28RA* ($P = 0,0008$), el *OMSR* ($P = 0,002$) y *IL22RA2* ($P = 0,003$).

En una fase de validación solo el SNP rs243324 de *SOCS1* se validó como factor de riesgo para EM en una cohorte independiente de 3.919 pacientes con EM y 4.003 controles [$P_{\text{Cochran-Mantel-Haenszel}} = 6 \times 10^{-5}$; OR (IC 95%) = 1,13 (1,07-1,20)]. Además, se observó que el alelo T del rs243324 estaba aumentado en pacientes con EM en las formas remitente recurrente y secundaria progresiva frente a los pacientes con la forma primaria progresiva [$P_{\text{Cochran-Mantel-Haenszel}} = 0,0096$; OR (IC 95%) = 1,24 (1,05-1,46)]. La asociación con *SOCS1* parece ser independiente del locus de riesgo *CLEC16A*.

2. Estudio de asociación a gran escala del gen *APOE*

Realizamos un meta-análisis de dos SNPs del gen *APOE* (rs429358 y rs7412) incluyendo datos procedentes de tres fuentes: datos de los GWAS, datos sistemáticamente recopilados de la literatura y datos de colecciones independientes que se han genotipado para este fin. Estos estudios incluyen los datos combinados de 13.913 casos y 15.831 controles [$P = 0,259$; OR (IC 95%) = 0,95 (0,88-1,04) y $P = 0,0569$; OR (IC 95%) = 1,07 (1,00-1,15),

para el rs429358 y rs7412, respectivamente]. A pesar de que nuestro estudio presenta potencia suficiente para detectar asociaciones a umbrales de significación del genoma completo, nuestros análisis no apoyan un rol del rs429358 o del rs7412 en la susceptibilidad a EM.

3. Follow-up del GWAS desarrollado por el IMSGC y estudio de otros genes candidatos

Se estudiaron los loci *PKN2* (rs305217), *GTF2B* (rs7538427), *EPHA4* (rs1517440), *YTHDF3* (rs12115114), *ANKFN1* (rs17758761) y *PTPRM* (rs479857) que no alcanzaron el umbral de significación en un estudio de validación de las señales de significación del primer GWAS; el gen *TMEM39A* (rs1132200), descrito como gen de susceptibilidad en el estudio de validación mencionado; el gen *CBLB* (rs9657904), que alcanzó significación a nivel de GWAS en población de Italia; el gen *IL12B* (rs6887695, rs10045431), que es un loci de susceptibilidad compartido por diversas enfermedades autoinmunes; y el gen *CNR1* (rs1049353), que mostraba una asociación poco concluyente con EM.

El análisis conjunto en 2.863 casos y 2.930 controles corroboró el efecto de tres loci en la predisposición a EM: *TMEM39A* [P = 0,001; OR (IC 95%) = 0,84 (0,75-0,93)], *IL12B* [P = 0,03; OR (IC 95 %) = 1,09 (1,01-1,17)] y *CBLB* [P = 0,01; OR (IC 95%) = 0,89 (0,81-0,97)].

4. Estudio de cinco genes sugerentes de asociación tras GWAS

En este estudio nos propusimos validar en una base de datos independiente, los cinco SNPs [*MANBA* (rs228614), *CXCR5* (rs630923), *SOX8* (rs2744148), *RPS6KB1* (rs2744148) y *ZBTB46* (rs6062314)] que habían mostrado una evidencia de asociación fuerte con el riesgo a padecer EM pero por debajo de los niveles de significación del barrido genómico más reciente. En un total de 8.499 casos, 8.765 controles y 958 tríos descendientes de europeos, todos los SNPs estudiados mostraron evidencia estadística significativa de asociación con riesgo a padecer EM: rs228614 [P = $2,4 \times 10^{-6}$; OR (IC 95%) = 0,91 (0,87-0,94)], rs630923 [P = $1,2 \times 10^{-4}$; OR (IC 95%) = 0,89 (0,84-0,95)], rs2744148 [P = $1,8 \times 10^{-6}$; OR (IC 95%) = 1,14 (1,08-1,20)], rs180515 [P = $5,2 \times 10^{-7}$; OR (IC 95%) = 1,12 (1,07-1,17)], rs6062314 [P = $4,3 \times 10^{-3}$; OR (IC 95%) = 0,90 (0,83-0,97)]. Al combinar nuestros datos con los resultados del GWAS previo mediante un meta-análisis, se reforzó la evidencia de asociación,

sobrepasando el umbral de significación del análisis de genoma completo ($P < 5 \times 10^{-8}$) en todos los casos.

5. Mapeo fino y análisis funcional del gen de riesgo CD6

En este estudio, se llevó a cabo un *tagging* de SNPs del locus *CD6*, tras el cual se seleccionaron 13 SNPs, además se incluyó para su estudio el SNP rs650258, asociado en el último GWAS de EM. Este conjunto de SNPs se estudió en una colección de 823 casos y 813 controles. La validación de los seis SNPs más fuertemente asociados se llevó a cabo en una colección independiente de 2.265 casos y 2.600 controles. El haplotipo TC marcado por dos SNPs no sinónimos [rs11230563 (R225W) y rs2074225 (A257V)] en el segundo dominio SRCR se identificó asociado con la susceptibilidad a padecer EM ($P_{\text{Permutada}} = 1 \times 10^{-4}$). También se estudió el efecto de este haplotipo en la expresión de CD6 en superficie y en la secreción de citoquinas. El análisis mostró patrones de expresión de CD6 diferentes en los distintos subtipos celulares: linfocitos T (LTs) $CD4^{+}$ naïve, $P = 0,0001$; LTs $CD8^{+}$ naïve, $P < 0,0001$; LTs de memoria central $CD4^{+}$ y $CD8^{+}$, $P = 0,01$ y $0,05$, respectivamente; y células NKT, $P = 0,02$; siendo el haplotipo protector TC el que mostraba la expresión más alta de CD6. Sin embargo, no se observaron cambios significativos en células NK, LTs de memoria efectores y de memoria efectores diferenciados terminalmente.

6. Influencia de la delección en *LILRA3* en el riesgo a desarrollar EM

En este estudio se incluyeron 3.200 casos y 3.069 controles procedentes de 5 colecciones de datos previamente publicadas y de 3 genotipadas para este fin. No se observó ninguna evidencia de asociación significativa de la delección *LILRA3* [*carriers* de *LILRA3*: $P = 0,25$, OR (IC 95%)= 1,07 (0,95-1,19)], ni después de realizar una estratificación adicional por género y el alelo de riesgo *HLA-DRB1*15:01*.

Conclusiones

Los avances llevados a cabo mediante la realización del presente trabajo permiten extraer las siguientes conclusiones:

1. El SNP rs243324 del gen *SOCS1* es un marcador de riesgo a EM y esta asociación es independiente del gen de riesgo *CLEC16A* previamente descrito.
2. Los SNPs rs7412 y rs429358, localizados en el gen *APOE*, no parecen modificar la susceptibilidad a EM en poblaciones de ascendencia europea, como muestra nuestro estudio que incluye 29.744 y 28.460 sujetos, respectivamente, y combina los resultados de genotipos *de novo* de múltiples poblaciones, datos imputados de GWAS y evidencia sistemáticamente cotejada de la literatura.
3. Los genes *IL12B*, gen candidato asociado con diversas enfermedades inmuno-mediadas; y *CBLB*, gen asociado inicialmente solo en la población italiana, se asocian al riesgo a presentar EM.
4. El gen *TMEM39A*, identificado previamente en un estudio de replicación del primer GWAS en EM y los genes *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* y *ZBTB46* previamente descritos al borde de la significación tras análisis por GWAS, se confirman en este estudio como loci de susceptibilidad a EM.
5. El haplotipo CT formado por los SNPs rs11230563 y rs2074225 del gen *CD6*, confiere riesgo a EM y modifica significativamente la expresión de CD6 en células CD4⁺ y CD8⁺.
6. La delección en el gen *LILRA3* no está asociada a EM por sí misma ni cuando se tienen en cuenta adicionalmente factores como el alelo *HLA*DRB1*15:01* o el género.

Introduction

Multiple sclerosis (MS) is a chronic inflammation of the central nervous system characterized by demyelization, multifocal inflammation, reactive gliosis, oligodendrocyte death and axonal loss. The etiology of MS is unknown, but it is considered a multifactorial disorder that is likely to occur in genetically susceptible subjects under the interaction of one or more environmental factors.

The *HLA-DRB1*15:01* allele is the main genetic factor associated with MS so far, but there are multiple non-HLA genes with a lower effect that also confer risk to develop this disease. Linkage and association studies have been performed to identify non-HLA risk loci, at first through candidate-gene approaches and more recently through genome wide association studies (GWAS). However, a percentage of heritability still remains uncovered (“missing heritability”) probably as a consequence of multiple factors, some of them studied in this work.

Objectives

1. To search for MS susceptibility genes by performing a screening of 386 polymorphisms located in genes coding for cytokines, cytokine receptors, signal transducers and ionotropic receptors and transporters of glutamate and to validate the most statistically significant signals in an independent cohort.
2. To perform a large scale association study of two polymorphisms in the candidate gene *APOE* using data from different origins.
3. To analyze the influence on the risk of developing MS of genes previously associated to other immune-mediated diseases or those explored only locally.
4. To validate polymorphisms previously identified through GWAS as suggestive risk factors for MS and association signals that did not reach genome-wide significance.
5. To identify the variant or haplotype responsible of the association of the *CD6* gene with MS susceptibility and to study its functional role.

6. To search for genetic susceptibility factors not included in GWAS that could help to explain the “missing heritability” in MS: the *LILRA3* gene.

Materials and methods

1. Cases and control subjects

This work includes information of a maximum of 842 MS patients and 832 controls of Caucasian origin collected in the Hospital Clínico San Carlos.

2. Manipulation of the samples and genotyping method

DNA isolation was performed by a salting out method from 10 ml of peripheral blood. Allele discrimination of single nucleotide polymorphisms (SNPs) was performed with Taqman assays using an ABI PRISM 7900HT analyzer.

3. Design of multicentric studies

The present work has been carried out in collaboration with different national and international research groups. For this, a maximum of 9457 cases of MS and 10 691 controls of Caucasian origin from different populations were genotyped. In some cases, systematic reviews were performed in the PubMed dataset of NCBI with subsequent data abstraction. Also, genotype data from the three GWAS datasets were obtained after request, depurated with PLINK v.1.07 and used to impute new data through IMPUTE v.2.0.

4. Selection of SNPs

We screened 368 SNPs in 55 genes or gene clusters coding for cytokines, cytokine receptors, suppressors of cytokine signaling, complement factors and glutamate receptors for association with MS. In addition, a total of thirteen *CD6* haplotype-tagging SNPs were selected for the fine mapping of the locus including nine intronic SNPs, three force-included non-synonymous SNPs and the previously reported rs17824933. Haplotype-tagging SNPs were selected using the multimarker tagger algorithm implemented in HapMap on the CEU cohort (r^2 cutoff 0.8; minor allele frequency 0.2; HapMap Release #19).

Other SNPs were selected based on previous GWAS, on their association with other immune-mediated diseases or based on previous gene-candidate studies on MS.

5. Genotyping of SNPs

A customized Bead Array Matrix SNP panel was manufactured for screening of 384 SNPs at Illumina Customer Service. In other cases, genotyping of SNPs was performed either in the participant center or in specialized genotyping centers such as the Spanish National Genotyping Center and the Max-Planck-Institute for Molecular Genetics, according either to the particular needs of each study or the decision of the PI in each localization.

These studies have a multicentric approach; thus, different quality assessments were checked to grant a coherent genotyping of the sample within each center and within the different centers.

6. Strategy for *LILRA3* genotyping

Both the complete and the deleted forms of *LILRA3* in homozygous and heterozygous combinations were detected in a single-tube PCR with three primers: two forward and one reverse, which were then distinguished by their different electrophoretic mobilities in regular agarose gels.

7. Samples for Fluorescence-activated Cell Sorting (FACS) and ELISA

10 ml of fresh lithium heparinized blood were obtained by venipuncture from twenty-seven MS patients and from twelve healthy donors. Peripheral blood mononuclear cells were purified, counted and immediately used or cryopreserved in the presence of dimethyl sulphoxide 10% (v/v) (DMSO) and 20% fetal bovine serum, and stored at -196°C until further use. When using frozen samples, PMBCs were thawed immediately before use.

8. Flow Cytometry

Anti-CD3 and anti-CD56 monoclonal antibodies were used to identify T, NK, and NKT cell subsets. Subsequently, anti-CD45RA, anti-CD27 and anti-CD28 monoclonal antibodies were used to distinguish naïve, central memory cells and effector memory cells, as well as the terminally differentiated effector memory cells. For staining, PBMCs were thawed, washed twice, and incubated 45 min with the antibody cocktails. Cells were subsequently washed, re-suspended in 300 μl FACS buffer and analyzed on a FACS Canto flow cytometer. FACS Diva software v.6.1.3 was used to analyze the raw data.

9. Cell culture, proliferation and ELISA assays for cytokine measurement

PBMCs (10^6 cells) from each donor were suspended in 500 μ l of RPMI containing 5% FBS, aliquoted into a 48-well plate and cultured under three different conditions: unstimulated, stimulated with anti-CD3 with or without anti-CD6 mAb. To determine cell proliferation, PBMC from healthy donors were labeled with 2 μ M eFluor670. To stop labeling, standard medium was added and incubated on ice for five minutes. Labeled cells were then transferred into 48-well plates with the respective stimuli for further cell culture.

Cells were counterstained with specific lineage markers to assess proliferation at day 0 and 3 by assessing the percentage of cells that had undergone division determined by flow cytometry. The supernatants from the culture were collected on the third day and quantified for IL-17 and IFN- γ using Ready-Set-Go ELISA kits. The cytokine concentration (pg/ml) was calculated using the standard curves generated using the respective standards.

10. Analysis of data

Statistical power was calculated using either the Genetic Power Calculator at <http://pngu.mgh.harvard.edu/~purcell/gpc/cc2.html> or the CATS power calculator at <http://www.sph.umich.edu/csg/abecasis/CaTS>. Genotypic, allelic and haplotypic comparisons in the case-control studies were performed using χ^2 -tests or the Fisher's exact test when values were below 5, with the Stalcalc software (EpiInfo v.6). Haplotype analyses were performed with Haploview or PLINK software. Interaction between independent loci and gender influence were performed by direct counting in stratified tables. Meta-analyses were performed with Review Manager 5.0, HardyWeinberg v.1.4 and rmeta v.2.16 statistical packages in R language, and with PLINK and METAL softwares. Logistic regression was performed in PLINK. ANOVA and two-tailed Student's-t test were used to calculate differences in surface expression, while differences in proliferation and cytokine production among the different haplotypes were assessed by the non-parametric Mann-Whitney U test using Graphpad Prism software v.5.

Results and discussion

1. Screening of candidate genes and validation of significant signals

Top-scoring SNPs in the screening phase were found within or nearby the genes coding for SOCS-1 ($P = 0.0005$), IL28RA ($P = 0.0008$), OMSR ($P = 0.002$), and IL22RA2 ($P = 0.003$). The *SOCS1* rs243324 variant was validated as a risk factor for MS in an independent cohort of 3919 MS patients and 4003 controls [$P_{\text{Cochran-Mantel-Haenszel}} = 6 \times 10^{-5}$; OR (95% CI) = 1.13 (1.07-1.20). In addition, the rs243324*T allele was consistently increased in relapsing-remitting/secondary progressive versus primary-progressive MS patients in each one of the six data sets used in this study [$P_{\text{CMH}} = 0.0096$; OR (95% CI) = 1.24 (1.05-1.46)]. The association with *SOCS1* seems to be independent from the one described in the same locus, *CLEC16A*.

2. Large scale association study of the gene APOE

We performed a meta-analysis of the most commonly studied *APOE* SNPs (rs7429358 and rs7412) in relation with MS susceptibility. To this end, we combined *de novo* genotyping results from multiple populations, imputed GWAS data, and systematically collated evidence from the literature. This study included meta-analyses of the combined data across 13 913 MS cases and 15 831 controls ($P = 0.295$; OR (95% CI) = 0.95 (0.88-1.04), and $P = 0.0569$; OR (95% CI) = 1.07 (1.00-1.15), for rs429358 and rs7412, respectively]. Given the large sample size of our analyses, it is unlikely that the two *APOE* missense SNPs studied here exert any relevant effect on MS susceptibility.

3. Follow-up to the GWAS performed by the IMSGC and study of other candidate genes

Replication studies were extended to the following polymorphisms: *PKN2* (rs305217), *GTF2B* (rs7538427), *EPHA4* (rs1517440), *YTHDF3* (rs12115114), *ANKFN1* (rs17758761) and *PTPRM* (rs4798571), which did not reach the threshold of significance in a follow-up of the first GWAS conducted in MS; *TMEM39A* (rs1132200), which appeared as a newly identified susceptibility gene in the same study; a gene previously reaching GWAS significance only in Italy, *CBLB* (rs9657904); *IL12B* (rs6887695, rs10045431), a susceptibility gene shared by diverse autoimmune diseases and, finally, another gene showing inconclusive association with multiple sclerosis, *CNRI* (rs1049353).

Pooled analysis of 2863 cases and 2930 controls corroborated the effect on MS predisposition of three genes: *TMEM39A* [rs1132200: $P_{M-H} = 0.001$; OR (95% CI) = 0.84 (0.75-0.93)], *IL12B* [rs6887695: $P_{M-H} = 0.03$; OR (95% CI) = 1.09 (1.01-1.17)] and *CBLB* [rs9657904: $P_{M-H} = 0.01$; OR (95% CI) = 0.89 (0.81-0.97)].

4. Study of 5 genes suggesting association with MS beyond the original GWAS

We tested five SNPs: rs228614 (*MANBA*), rs630923 (*CXCR5*), rs2744148 (*SOX8*), rs180515 (*RPS6KB1*), and rs6062314 (*ZBTB46*), for association with MS risk in a total of 8499 MS cases, 8765 unrelated control subjects and 958 trios of European descent. In addition, we assessed the overall evidence for association by combining these newly generated data with the results from the original GWAS by performing a meta-analysis. All five tested SNPs showed consistent and statistically significant evidence for association with MS in our validation data sets: rs228614 [$P = 2.4 \times 10^{-6}$; OR (95% CI) = 0.91 (0.87-0.94)], rs630923 [$P = 1.2 \times 10^{-4}$; OR (95% CI) = 0.89 (0.84-0.95)], rs2744148 [$P = 1.8 \times 10^{-6}$; OR (95% CI) = 1.14 (1.08-1.20)], rs180515 [$P = 5.2 \times 10^{-7}$; OR (95% CI) = 1.12 (1.07-1.17)], rs6062314 [$P = 4.3 \times 10^{-3}$; OR (95% CI) = 0.90 (0.83-0.97)]. Combining our data with the results from the previous GWAS by meta-analysis, the evidence for association was further strengthened, surpassing the threshold for genome-wide significance ($P < 5 \times 10^{-8}$) in each case.

5. Fine mapping and functional analysis of the MS risk factor *CD6*

In this study, we performed a haptag SNP screen of the *CD6* locus using a total of thirteen tagging SNPs, of which three were non-synonymous SNPs, and we replicated the association recently reported by GWAS of the SNP rs650258 in a collection of 814 controls and 823 cases. Validation of the six most strongly associated SNPs was performed in an independent collection of 2265 MS patients and 2600 healthy controls. We studied the association of the haplotypes composed of two non-synonymous SNPs [rs11230563 (R225W) and rs2074225 (A257V)] in the 2(nd) SRCR domain with susceptibility to MS ($P_{\max(T) \text{ permutation}} = 1 \times 10^{-4}$). A protective haplotype (RA) was evidenced and the effect of this haplotype on CD6 surface expression and cytokine secretion was also tested. Our analysis showed significantly different CD6 expression patterns in the distinct cell subsets, i.e. - CD4⁺ naïve cells, $P = 0.0001$; CD8⁺ naïve cells, $P < 0.0001$; CD4⁺ and CD8⁺ central

memory cells, $P = 0.01$ and 0.05 , respectively; and natural killer T (NKT) cells, $P = 0.02$; with the protective haplotype (RA) showing higher expression of CD6. However, no significant changes were observed in NK cells, effector memory and terminally differentiated effector memory T cells.

6. Influence of the *LILRA3* deletion on the risk of developing MS

This meta-analysis involved 3200 MS patients and 3069 matched healthy controls. With this aim, we collected and genotyped 3531 subjects hitherto no studied for their *LILRA3* status, and extracted genotype data for 2828 subjects from previous candidate-gene assessments. Our study provided no evidence of a significant association of the *LILRA3* deletion [carriers of *LILRA3* deletion: $p = 0.25$, OR (95% CI) = 1.07 (0.95-1.19)] with MS risk, after stratification either by gender or by the *HLA-DRB1*15:01* risk allele.

Conclusions

1. The *SOCS1* rs243324 variant is a risk factor for MS and this association is independent from the *CLEC16A* mapping to the same locus
2. The SNPs rs4029358 and rs7412, located in the *APOE* gene, do not seem to modify the risk for MS in populations of European descent, as our study with 29 744 cases and 28 460 controls combining *de novo* genotyping results from multiple populations, imputed GWAS data, and evidence systematically reviewed from the literature.
3. The *IL12B* gene, a candidate gene associated with several immune-mediated diseases, and the *CBLB* gene, which has been previously associated with MS risk only in the Italian population, are associated with the risk of developing MS.
4. The *TMEM39A* gene, previously identified in a follow up study the first GWAS, and the *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* and *ZBTB46* genes previously showing sub-genome-wide significant associations in a GWAS, are confirmed as MS susceptibility loci after our studies.

5. The haplotype composed of two non-synonymous SNPs, rs11230563 and rs2074225 of the *CD6* gene, confers risk of for MS and significantly modifies the expression of CD6 in CD4⁺ and CD8⁺ cells.
6. The *LILRA3* gene does not seem to be underlying MS pathogenesis even when other factors such as *HLA*DRB1*15:01* allele or gender are considered.

Introducción

1. Definición y epidemiología

La esclerosis múltiple (EM, OMIM 126200) es una enfermedad del sistema nervioso central (SNC) caracterizada por desmielinización, inflamación multifocal, gliosis reactiva y pérdida oligodendrocitaria y axonal ¹. La EM es una de las causas más comunes de discapacidad neurológica en adultos jóvenes, afectando a más de 2,3 millones de personas en todo el mundo ².

En Europa, la tasa de prevalencia estimada de la EM es de 108 casos por cada 100.000 individuos, con tasas más altas en países del norte y un *ratio* hombre:mujer de alrededor de 2 (Figura 1). La frecuencia de la EM es baja en la infancia e incrementa después de los 18 años, alcanzando un pico entre los 20 y los 40 años (media de edad de 30 años) y suele aparecer de 2 a 5 años antes en las mujeres que en los hombres ³.

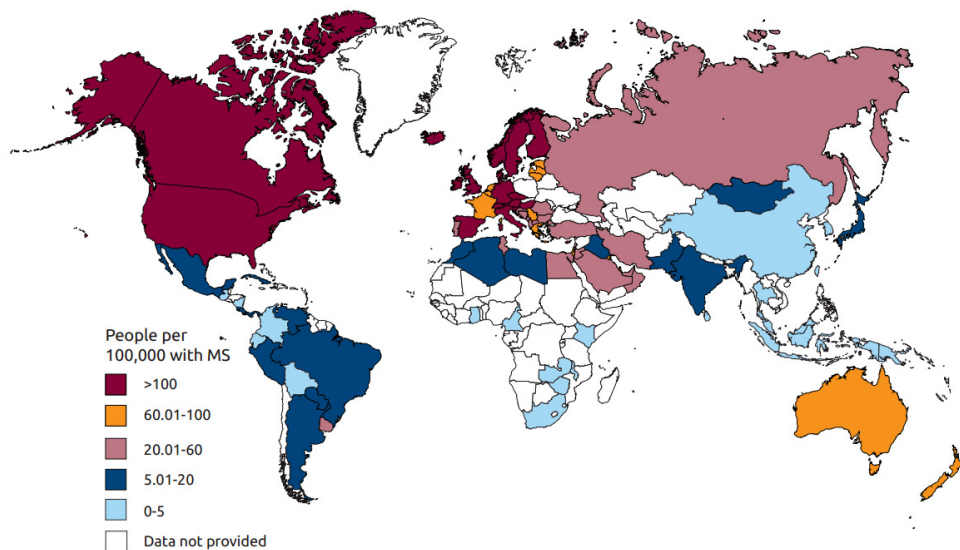


Figura 1. Prevalencia mundial de la EM (2013). Figura extraída del Atlas de Esclerosis Múltiple del 2013 ².

2. Formas clínicas

El curso clínico de la EM muestra heterogeneidad entre pacientes y también dentro del mismo paciente ⁴. La siguiente categorización del curso clínico de la EM se ha reconocido por consenso internacional para reducir la confusión en la terminología (Figura 2) ^{5,6}:

- **EM Remitente recurrente (EMRR).** Se caracteriza por un abrupto comienzo de los síntomas y episodios agudos de empeoramiento (exacerbaciones o brotes) con remisión total o parcial. Es la forma clínica que presentan el 85% de los pacientes con EM.
- **EM Secundaria progresiva (EMSP).** Se caracteriza por la persistencia de los signos de disfunción del SNC que pueden desarrollarse después de un brote. Entre un 30 y un 50% de los pacientes que sufren inicialmente EMRR, desarrollan la forma secundaria progresiva.
- **EM Primaria progresiva (EMPP).** Forma crónica caracterizada por la existencia de una progresión gradual de la discapacidad desde el inicio de la enfermedad, sin brotes sobrepuestos. Aparece en el 10% de los pacientes afectados.
- **EM Progresiva recurrente (EMPR).** Los pacientes experimentan una progresión gradual de la discapacidad desde el inicio de la enfermedad, más tarde acompañada por uno o más brotes. Este patrón clínico afecta aproximadamente al 5% de los pacientes con EM.

A menudo, debido al diseño transversal e histórico de muchos estudios epidemiológicos en la EM, el curso clínico de la enfermedad se categoriza de forma más simple en EMRR, EMSP y EMPP; basándose en la mayor prevalencia de casos ⁷.

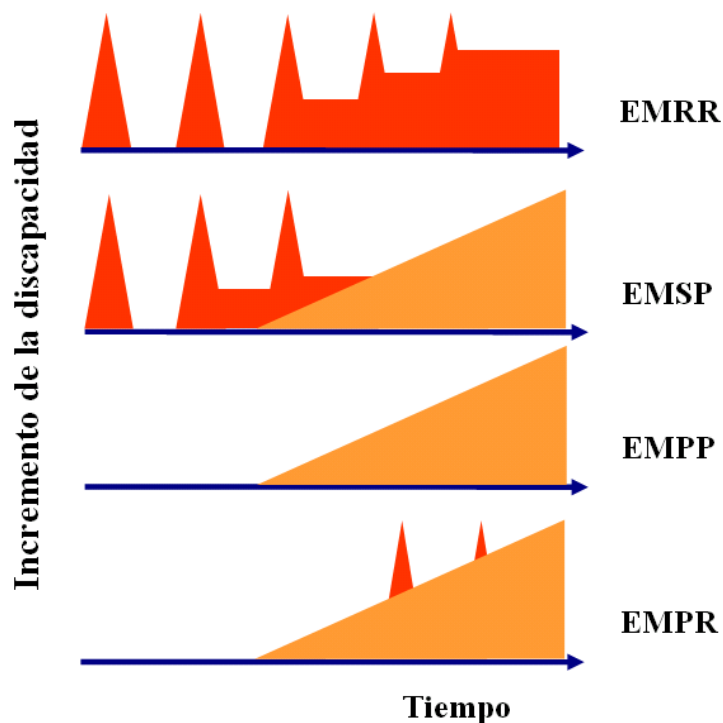


Figura 2. Cursos clínicos de la EM. Áreas rojas: brote; áreas naranjas: progresión de la enfermedad. Figura adaptada de Kamm *et al.*⁸

3. Diagnóstico

La EM es una enfermedad neurodegenerativa crónica del SNC en la que brotes de inflamación inmunomediada resultan en desmielinización, pérdida de oligodendrocitos y degeneración axonal. Esto a su vez, resulta en una multiplicidad de déficits, con implicación de los sistemas motor, sensorial, visual y autonómico; tales como parálisis, pérdida sensorial, debilidad de las extremidades, fatiga, calambres, incontinencia y defectos cognitivos¹.

El principio del diagnóstico es establecer que la actividad de la enfermedad ha afectado a más de una parte del SNC y en más de una ocasión mediante evidencias clínicas y estudios complementarios⁹. Los estudios paraclínicos que se pueden usar para estudiar una sospecha de EM incluyen:

- **Imágenes de resonancia magnética (RM).** Se pueden usar para establecer la diseminación en el espacio y en el tiempo de las lesiones en el SNC. Las imágenes de RM muestran anormalidades focales en la materia blanca en más del 95% de los pacientes de EM. Sin embargo, su sola presencia no permite el diagnóstico de EM ⁹.
- **Análisis del líquido cefalorraquídeo (LCR).** La presencia de bandas oligoclonales de inmunoglobulina G en el LCR por isoelectroenfoque sugiere síntesis intratecal de inmunoglobulinas. Se observa en aproximadamente el 90% de los pacientes ¹⁰.
- **Potenciales visuales evocados (PVE).** Son respuestas eléctricas corticales totales a estímulos sensoriales periféricos que se pueden usar para localizar sitios de patología y medir la velocidad de conducción a lo largo de las vías sensoriales ¹¹.

Los criterios diagnósticos más ampliamente usados durante las dos últimas décadas del siglo XX fueron los criterios establecidos por Poser *et al.* ¹² Dichos criterios permitían la clasificación de los pacientes con EM de acuerdo al número de brotes, las evidencias clínicas y paraclínicas, la presencia de bandas oligoclonales de inmunoglobulina G y/o un incremento del índice de IgG. Sin embargo, en el 2001, McDonald *et al.* ¹³ revisaron los criterios de Poser, con el propósito de integrar las imágenes de RM en el diagnóstico de la EM (Tabla 1).

Tabla 1. Criterios diagnóstico de McDonald *et al.* ¹³

Clínico Brotos	Objetivo MRI	Requisitos adicionales para el diagnóstico
2 o más	2 o más	Ninguno. La evidencia clínica es suficiente. Otra evidencia es deseable pero deber ser consistente con la EM
2 o más	1	Diseminación en el espacio demostrada por RM o 2 o más lesiones en la RM consistente con EM más LCR positivo o se espera otro brote clínico
1	2 o más	Diseminación en el tiempo demostrada por la RM o un segundo brote clínico
1	1	Diseminación en el espacio demostrada por la RM o 2 o más lesiones en la RM consistentes con EM más LCR positivo Y diseminación en el tiempo demostrada por RM o un segundo brote clínico
0 (Progresión desde el inicio)	1 o más	Progresión de la enfermedad por 1 año (retrospectivo o prospectivo) y 2 de los siguientes 3 hallazgos: <ul style="list-style-type: none">▪ RM del cerebro positiva (9 lesiones en T2 ó 4 o más lesiones en T2 con PVE positivo)▪ RM de la columna positiva (2 o más lesiones focales en T2)▪ LCR positivo

RM: resonancia magnética; LCR: líquido cefalorraquídeo; PVE: potenciales visuales evocados.

4. Inmunopatogénesis de la esclerosis múltiple

La alteración de la regulación de las respuestas inflamatorias y de la tolerancia inmunitaria se consideran elementos claves en la respuesta inmunitaria autorreactiva que tiene lugar en la EM ¹⁴. Se cree que la EM es un trastorno autoinmune, pero no se ha podido determinar una única especificidad antigénica que provoque esta respuesta, a pesar de los considerables esfuerzos que se han realizado al respecto ^{15, 16}. La hipótesis de que la EM es una enfermedad autoinmune proviene principalmente de las similitudes observadas entre la

EM y la encefalitis autoinmune experimental (EAE). La EAE se puede inducir en una variedad de especies animales, incluyendo primates no humanos, por inmunización con proteínas (o péptidos) derivados de la mielina, como la proteína proteolipídica, la glicoproteína oligodendrocitaria de la mielina y la proteína básica de la mielina¹⁷. La EM y la EAE comparten características clínicas, histológicas, inmunológicas y genéticas comunes; por lo tanto, la EAE se considera un modelo animal relevante de esta enfermedad humana.

4.1. Procesos inmunológicos en la esclerosis múltiple

Se han postulado múltiples mecanismos de daño en la mielina mediados por el sistema inmunitario en la EM (Figura 3). Algunos de los procesos observados son:

- Cambios inmunológicos en la periferia

Los linfocitos T (LTs) tienen un papel esencial dirigiendo la compleja cascada de eventos en la EM, dentro de los que se incluye la inflamación crónica, la desmielinización primaria y el daño axonal¹⁴. La activación de LTs CD4⁺ en la periferia se considera un evento inmunitario importante en la EM. La forma en la que estos LTs autorreactivos se activan en la periferia en pacientes con EM es todavía incierta. Se han postulado como mecanismos potenciales, procesos tales como el mimetismo molecular, donde LTs generados contra epítomos no propios (antígenos virales o microbianos) tienen reacción cruzada con epítomos propios de la mielina de similar secuencia^{18, 19}; o una activación de LTs desencadenada por antígenos de mielina constitutivamente presentados en los nódulos linfáticos cervicales^{20, 21}.

- Migración al sistema nervioso central

Posteriormente, los LTs autorreactivos migran al interior del parénquima cerebral a través de un proceso secuencial mediado principalmente por moléculas de adhesión, quimiocinas y metaloproteinasas de matriz. Los pasos iniciales en la migración a través de la barrera hematoencefálica (BHE) implican una interacción entre las moléculas de adhesión expresadas tanto en células endoteliales como en células del sistema inmunitario. Una

molécula de adhesión clave implicada en la extravasación leucocitaria es la integrina $\alpha 4 \beta 1$ (también conocida VLA-4), que se expresa en la superficie de linfocitos activados e interactúa con las moléculas de adhesión celular vascular (VCAM-1, *vascular cell adhesion molecule 1*) expresada en células endoteliales capilares ^{16, 22}.

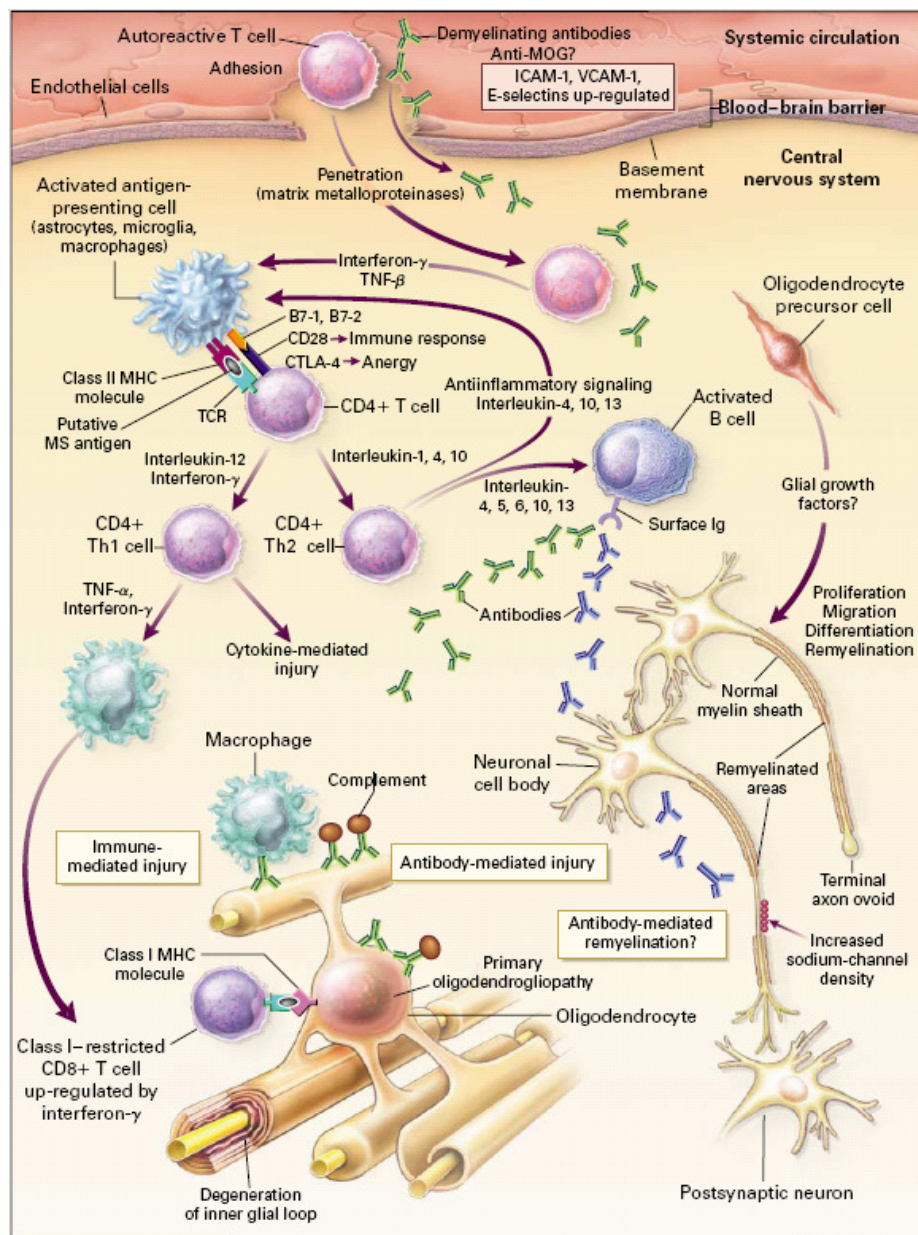


Figura 3. Immunopatogénesis de la EM. Figura extraída de Noseworthy *et al.* ¹

- Interacción de los LTs autorreactivos con el antígeno diana

Dentro del SNC, los LTs CD4⁺ específicos de la mielina son reactivados *in situ* por antígenos de la mielina presentados por moléculas de HLA de clase II presentes en la superficie de células presentadoras de antígeno como los macrófagos y la microglía, en conjunción con otras proteínas accesorias. La reactivación desencadena la liberación de citoquinas proinflamatorias (IFN- γ , IL-23, TNF- α , linfotoxina, etc.) y quimioquinas, que alteran aún más la BHE y estimulan la quimiotaxis, lo que resulta en una segunda gran ola de reclutamiento de células inflamatorias en el SNC, incluyendo monocitos, LTs CD8⁺, linfocitos B (LBs) y mastocitos de sangre periférica, que contribuyen a la formación de la lesión inflamatoria. Además, esta cascada de eventos inicial puede dar lugar a la liberación y alteración de otros antígenos tisulares, a la activación de linfocitos específicos frente a esos otros antígenos y a la exacerbación de la enfermedad. Este fenómeno conocido como propagación del epítipo probablemente desempeñe un papel importante en la progresión y la perpetuación de la respuesta autoinmune.

El daño en la mielina provoca que los axones ya no sean capaces de transmitir potenciales de acción de forma eficiente dentro del SNC. Esta disminución o bloqueo en los potenciales de acción provoca los síntomas neurológicos característicos de la enfermedad. Además, los segmentos de axón expuestos pueden ser susceptibles de daño adicional por mediadores solubles, tales como citoquinas, quimiocinas, factores del complemento y proteasas; que provocan un daño axonal irreversible. Hay varios posibles mecanismos de reparación de la vaina de mielina, entre los que se incluyen: la resolución de la respuesta inflamatoria seguida de remielinización espontánea, la redistribución de los canales de sodio de los nodos de Ranvier para cubrir los segmentos desnudos del axón y restaurar la conducción; y la remielinización resultante de la proliferación, migración y diferenciación de células precursoras de oligodendrocitos residentes ¹.

4.2. Componentes celulares de la respuesta inmunitaria en la esclerosis múltiple

4.2.1. Linfocitos T CD4⁺

Durante las pasadas décadas, las investigaciones en la EM se centraron principalmente en el papel de los LTs CD4⁺ en la patogénesis de la enfermedad. La noción de que la EM es ante todo una enfermedad mediada por LTs CD4⁺ proviene de los estudios realizados en la EAE, donde los LTs CD4⁺ específicos de la mielina son los que conducen a la enfermedad y al hecho de que ésta se puede transferir adoptivamente a animales no afectados mediante la inyección de estos LTs reactivos de mielina.

Después de un proceso de activación, los LTs vírgenes se diferencian en varias poblaciones de LTs con diferentes funciones efectoras. El papel de los diferentes subgrupos de LTs CD4⁺ ha sido objeto de investigación en EM e inicialmente se postuló que una alteración de la regulación del balance entre la respuesta inmunitaria Th1 y Th2 podría estar implicada en la patogénesis de la EM ^{23, 24}. Sin embargo, actualmente algunos estudios han mostrado también un papel de los linfocitos Th17. Por ejemplo, los estudios en la EAE han demostrado que la supresión de los linfocitos Th17 se asocia con una reducción de la severidad de la enfermedad ^{25, 26}. También se ha observado que el *ratio* Th17/Th1 se asocia con la infiltración de LTs e inflamación en el parénquima cerebral ²⁷ y que las células que expresan IL-17 o IFN- γ cruzan eficientemente la BHE y se acumulan dentro del cerebro ²⁸. Además, se ha visto que existe un aumento de la expresión de IL-17 en sangre, LCR y tejido cerebral de pacientes con EM ²⁹, así como un aumento significativo de linfocitos Th17 en LCR de pacientes con EM en los brotes clínicos comparado con las fases de remisión. También existe una expresión más alta de marcadores de activación y moléculas coestimuladoras y de adhesión en clones de linfocitos Th17 que en clones de linfocitos Th1 ³⁰, lo que apunta a un alto potencial patogénico de estas células. Finalmente en estudios *in vitro* con fingolimod, uno de los tratamientos orales empleados en la EM, se ha demostrado que suprime la generación de linfocitos Th17 ³¹.

Distintos estudios indican que en la EM existe una disfunción o un deterioro de la maduración de los linfocitos T reguladores (LT_{REG}). Así, se ha observado que la delección selectiva de LT_{REG} causa enfermedad autoinmune espontánea en ratones, mientras que un incremento de la función de los LT_{REG} puede prevenir el desarrollo o aliviar los síntomas de

algunas formas de la EAE. Por otro lado, el desarrollo y función de los LT_{REG} está estrechamente relacionado con las CDs, que tienen un papel central en la activación y reactivación de las células encefalitogénicas en el SNC^{14, 32}.

4.2.2. Linfocitos T $CD8^+$

Hay evidencias que sugieren un papel importante de los LTs $CD8^+$ en la EM. Entre ellas, se incluye la presencia de un alto número de LTs $CD8^+$ en el infiltrado inflamatorio en las lesiones del SNC, llegando incluso a superar en número a los LTs $CD4^+$, según algunos estudios³³. Existe una correlación entre el número de LTs $CD8^+$ que infiltran la lesión y el daño axonal³⁴; los LTs $CD8^+$ infiltrantes están clonalmente expandidos y pueden persistir en el LCR por muchos años³⁵. Además, los LTs $CD8^+$ pueden promover la permeabilidad vascular del SNC³⁶ y la transferencia adoptiva de clones de LTs $CD8^+$ induce EAE, sugiriendo un papel de los LTs $CD8^+$ como células efectoras en la patogénesis de la EM³⁷. Finalmente, en estudios *in vitro*, se ha observado que LTs $CD8^+$ pueden inducir lesiones a neuritas y podrían ser responsables del daño axonal a través de péptidos presentados por moléculas de HLA de clase I³⁸.

4.2.3. Linfocitos B

Se considera que la inmunidad humoral desempeña un papel muy importante en la patogénesis de la EM. En primer lugar, existe una producción intratecal persistente de inmunoglobulinas oligoclonales en el LCR de los pacientes con EM, un hallazgo que actualmente forma parte de los criterios diagnósticos de la enfermedad¹⁰. Además, los LBs pueden participar directamente en el proceso de desmielinización al secretar anticuerpos patogénicos dirigidos contra oligodendrocitos, con o sin la presencia de complemento³⁹; y que los LBs aislados del LCR y las lesiones cerebrales de pacientes con EM están expandidos clonalmente^{40, 41}. Otro hallazgo importante es la presencia de agregados tipo folículo con centros germinales en las meninges de algunos pacientes con un curso progresivo de EM, lo que sugiere que las respuestas de los LBs tales como la proliferación, la maduración de la afinidad inducida por antígeno y la diferenciación en células plasmáticas productoras de anticuerpos se puede mantener localmente dentro del SNC y puede contribuir

al proceso patogénico ⁴². Por último, el tratamiento ablativo de LBs con rituximab, un anticuerpo monoclonal anti-CD20 que reduce eficientemente los LBs vírgenes y de memoria, disminuye las lesiones cerebrales inflamatorias y los brotes clínicos en la EM ⁴³, supuestamente por una disminución de la capacidad presentadora de antígenos y de la producción de citoquinas por los LBs.

4.2.4. Células del sistema inmunitario innato en la patogénesis de la EM

Las células del sistema inmunitario innato pueden actuar de forma beneficiosa o perjudicial en la EM. Por un lado, pueden prevenir la autoinmunidad al inducir la diferenciación de LTs en LT_{REG} y pueden estimular la secreción de factores de crecimiento neurotróficos. Por otro lado, pueden desempeñar un papel inmunopatogénico al promover la diferenciación hacia células Th1 y Th17, las cuales dirigen los eventos inflamatorios agudos asociados con los brotes en la EM (Figura 4).

Además, se piensa que la fase progresiva de la EM está mediada por células del sistema inmunitario innato, como refleja el fenotipo activado de estas células en la periferia, que podría ser responsable directo de los cambios neurodegenerativos en la EMSP ⁴⁴. Hasta ahora, no existen terapias específicas dirigidas frente a las células del sistema inmunitario innato en la EM. A medida que se defina mejor su papel en la EM, será posible diseñar terapias para transformar células inmunopatogénicas en células inmunorreguladoras.

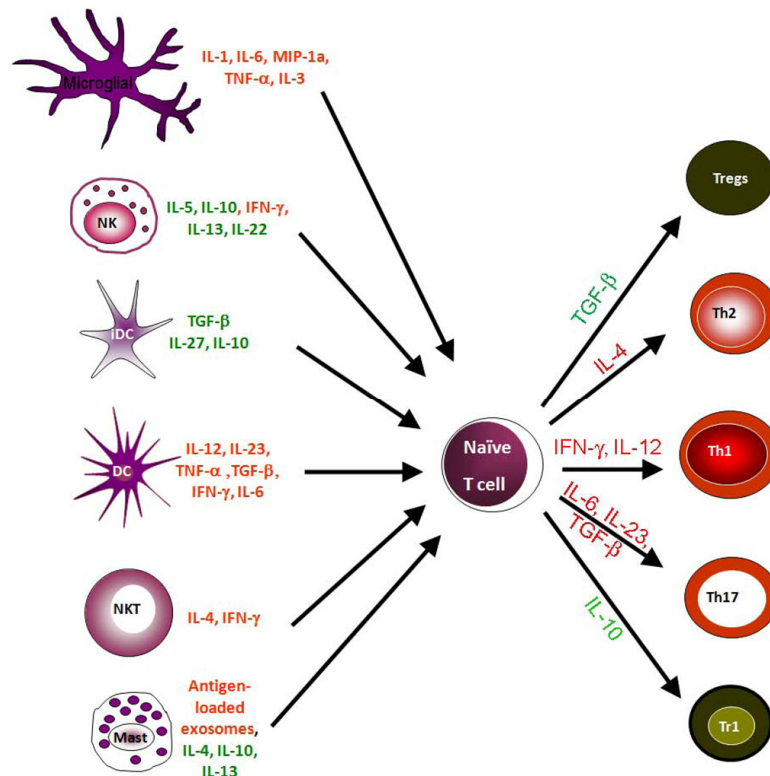


Figura 4. Citoquinas secretadas por células del sistema inmunitario innato en la EM y su papel en la diferenciación de LTs: LTs efectores o reguladores. Las citoquinas inflamatorias (en rojo) secretadas por células del sistema inmunitario innato conducen a la diferenciación de poblaciones de LTs efectores, tales como linfocitos Th1, Th2 y Th17, involucrados en los procesos de inflamación e inmunomodulación. Por otro lado, las citoquinas antiinflamatorias (en verde) secretadas por células del sistema inmunitario innato desempeñan un papel importante induciendo LT_{REG} que son capaces de mediar funciones tolerogénicas durante la enfermedad. Figura extraída de Gandhi *et al.*⁴⁵

5. Tratamiento

Debido a la considerable heterogeneidad clínica, genética, de imágenes de resonancia magnética y patológica de la EM, quizás más de un mecanismo contribuye al daño tisular. Esta posibilidad tiene implicaciones terapéuticas muy importantes porque se puede requerir más de un enfoque para tratar eficazmente esta enfermedad¹. El tratamiento actual de

primera línea para la EMRR y la EMSP consiste en terapias modificadoras de la enfermedad (TME), pues reducen las tasas de brotes y ralentizan la progresión de la discapacidad; e incluyen el interferón (IFN) β y el acetato de glatiramero ⁴⁶. Sin embargo, aunque estas terapias tienen un perfil de riesgo-beneficio favorable, son solo parcialmente efectivas ⁴⁷. Entre las posibles estrategias de tratamiento para pacientes que tienen una respuesta subóptima a una TME de primera línea, se incluye el cambio a otra TME de primera línea o a una TME de segunda línea ^{46, 48}. Además, en la actualidad hay nuevos fármacos en proceso de desarrollo clínico ⁴⁹.

A continuación se detallan algunos de los tratamientos de manera resumida:

5.1. Propiedades inmunomoduladoras

El IFN- β tiene diferentes formulaciones: IFN β -1a (*Avonex*® y *Rebif*®) e IFN β -1b (*Betaferón*® y *Extravia*®). Puede actuar a distintos niveles, tales como la inhibición de la activación y la coestimulación de LTs, la modulación de citoquinas antiinflamatorias y proinflamatorias, la disminución de la migración de LTs, la supresión de la diferenciación de linfocitos Th17 y la estimulación de la producción de NGF (*nerve growth factor*) en estadios tempranos de la enfermedad ¹.

El acetato de glatiramero (*Copaxone*®) es una mezcla de polipéptidos sintéticos que contiene los L-aminoácidos glutámico, alanina, lisina y tirosina. Puede promover la proliferación de citoquinas Th2, competir con la proteína básica de la mielina por la presentación en moléculas de HLA de clase II (inhibiendo la activación de LTs específicos de antígeno), alterar la función de los macrófagos e inducir LTs supresores específicos de antígeno ^{1, 50}.

5.2. Inhibición de la migración linfocitaria

El natalizumab (*Tysabri*®) es un anticuerpo monoclonal humanizado que reconoce la integrina $\alpha 4\beta 1$, inhibiendo la unión a su ligando (VCAM-1) e impidiendo la migración transendotelial de leucocitos a través de la BHE ⁵¹.

5.3. Propiedades inmunosupresoras

El fingolimod (*Gylenia*®) o FTY720 modula la activación de los receptores 1, 3-5 de la esfingosina 1 fosfato (S1P) y previene la salida de linfocitos desde el tejido linfoide secundario a los sitios de inflamación, reteniendo las células de memoria efectoras y los linfocitos Th17 diferencialmente. También, se piensa que podría promover la remielinización actuando en receptores 5 de S1P en oligodendrocitos^{52, 53}.

El alemtuzumab es un anticuerpo monoclonal que se une a CD52, un antígeno de superficie de función desconocida en linfocitos, monocitos y CDs. Induce una depleción sostenida de LTs y una depleción transitoria de LBs; incrementa los niveles de BAFF (*B-cell activating factor*) y de LT_{REG}; e incrementa la secreción de BDNF (*brain-derived neurotrophic factor*) por linfocitos⁵⁴.

5.4. Efecto sobre los linfocitos T

El daclizumab es un anticuerpo monoclonal humanizado del subtipo IgG1 que se une al epítipo Tac en la cadena α (CD25) del receptor de la IL-2, bloqueando así la formación del receptor de alta afinidad de la IL-2. Expande y activa células NK^{bright} inmunorreguladoras, bloquea la trans-presentación de IL-2 de CDs maduras a células T primadas y modula células del sistema inmunitario innato⁵⁵.

5.5. Efecto sobre los linfocitos B

El rituximab (*Mabthera*®), el ocrelizumab y el ofatumumab son anticuerpos monoclonales que se unen a diferentes epítomos del CD20, un antígeno de superficie expresado en LBs pero no en células plasmáticas. Causan una rápida depleción de los LBs⁵⁶.

6. Etiología de la esclerosis múltiple

La etiología de la EM es desconocida, sin embargo, se considera que es un trastorno multifactorial que se desencadena en individuos genéticamente susceptibles como

consecuencia de la interacción con uno o varios factores ambientales ^{1, 16}. Además, se piensa que efectos epistáticos y factores intrínsecos del respectivo individuo juegan un papel importante en la patogénesis de la EM (Figura 5).

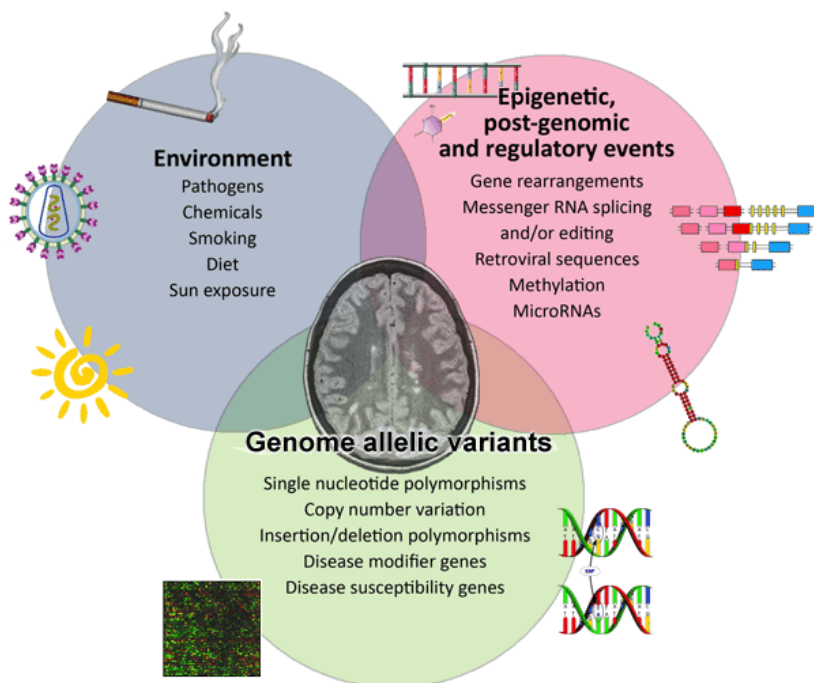


Figura 5. La esclerosis múltiple es una enfermedad compleja que no está causada por un único gen o factor ambiental. Presumiblemente, la susceptibilidad a la enfermedad y la representación fenotípica están influenciadas por la variación en numerosos genes en una interacción compleja con factores ambientales. (Imagen de Jorge Oksenberg, Universidad de California, San Francisco: <http://www.ucsf.edu/news/2011/08/10431/major-genetic-study-multiple-sclerosis-reveals-dna-hot-spots-disease>).

6.1. Factores de riesgo ambientales

Los estudios epidemiológicos ilustran claramente un papel destacado del ambiente en la susceptibilidad a padecer EM ⁵⁷. Una de las evidencias ambientales más sorprendentes es

el curioso gradiente latitudinal que existe en la EM, es decir, cuanto más alejada está una población del ecuador, más alta es la incidencia y prevalencia de EM, una correlación que, sin embargo, ha disminuido en las últimas décadas ^{4, 58, 59}. La latitud correlaciona con algunos factores ambientales, siendo uno de los más asociados, la exposición a la luz solar. El riesgo a EM se correlaciona inversamente con la exposición a la luz solar, y la vitamina D parece ser el agente protector responsable ⁶⁰.

Otro factor ambiental relacionado con la EM es el tabaco. Por ejemplo, se ha observado que el riesgo relativo de desarrollar EM es aproximadamente 1,5 veces mayor para fumadores comparado con no fumadores y que existe una asociación entre el hecho de fumar y el riesgo de conversión a EMSP, la progresión y la severidad de la enfermedad ⁶¹.

Parece ser que algunos microorganismos pueden contribuir al desarrollo de la EM en individuos genéticamente susceptibles. Sin embargo, no se ha identificado un único agente infeccioso causal para la EM y la existencia de tal agente es altamente improbable. No se sabe qué infección incrementa, disminuye o no influye en el riesgo a EM ⁶². Además, los estudios recientes en animales, sugieren que la microbiota comensal podría ser esencial en desencadenar procesos inmunitarios en ausencia de agentes patógenos ⁶³. Dentro de los agentes infecciosos que se han asociado a la EM cabe destacar el virus de Epstein-Barr, ya que se ha observado que el riesgo de la enfermedad es extremadamente bajo entre los individuos seronegativos, siendo los pacientes con EM seropositivos en más del 99% de los casos, contrastando con el 95% de la población general [OR (IC 95%) = 0,06 (0,02-0,13)]. Además, los individuos con una historia de mononucleosis infecciosa tienen un mayor riesgo de desarrollar EM frente a los que no, especialmente si la infección ocurre después de la adolescencia ⁶².

El tiempo de exposición a factores ambientales de riesgo juega un papel importante, siendo la infancia y la adolescencia las etapas de la vida más críticas. Esto se basa en los estudios migratorios entre las regiones de alto y de bajo riesgo. Las personas que migran a zonas con un riesgo alto antes de la adolescencia adquieren el riesgo de su nueva región, mientras que aquellas que lo hacen después de la adolescencia mantienen el riesgo de su región de origen ⁶⁴. Además, el riesgo a EM depende del mes de nacimiento. En el hemisferio norte, el riesgo de la EM es mayor para individuos nacidos en mayo y más bajo para aquellos nacidos en noviembre, lo que apoya a que las influencias del medio ambiente

en etapas tempranas de la vida, como infecciones estacionales por virus, podrían contribuir al riesgo a desarrollar EM ⁶⁵.

6.2. Factores de riesgo genéticos

6.2.1. El componente genético de la EM

Actualmente, la agregación familiar de la EM es ampliamente reconocida, ya desde 1988 cuando se publicaron estudios poblacionales sistemáticos para familias con miembros con EM ⁶⁶. Los estudios mostraron que familiares de primer, segundo y tercer grado de pacientes con EM eran más propensos a padecer la enfermedad con respecto a la población general y que los gemelos monocigóticos mostraban una alta concordancia en comparación con los gemelos dicigóticos ⁶⁷⁻⁷¹. En estudios llevados a cabo con enfermos de EM que habían sido adoptados, se observó un riesgo alto en la familia relacionada genéticamente que no fue observado en la familia adoptante ⁷². También se ha observado que la frecuencia de que ambos cónyuges estén afectados por la enfermedad no es mayor que la esperada por azar, pero el riesgo de que su descendencia esté afectada es mayor que si está afectado solo un progenitor ⁷³. Además, el riesgo para medios hermanos es aproximadamente la mitad del riesgo para hermanos completos, independientemente de si fueron criados juntos o no ^{74, 75}. Estos estudios muestran que los factores genéticos determinan la agrupación familiar y la susceptibilidad individual observada en la EM (Figura 6).

La EM es un desorden autoinmune que puede definirse como una enfermedad genéticamente compleja que no sigue una herencia mendeliana. Es difícil identificar variantes de susceptibilidad genética en enfermedades complejas como la EM porque los efectos atribuibles a un único gen son modestos y porque la variación genética en la población humana total es amplia y existen numerosas variantes genéticas que influyen en mayor o menor medida en el riesgo de padecer la enfermedad. Así, por ejemplo, la región HLA ha sido la más fuertemente asociada a la EM, pero diversos genes fuera de esta región también parecen contribuir al riesgo de desarrollo de esta enfermedad aunque con un efecto menor ⁷⁶.

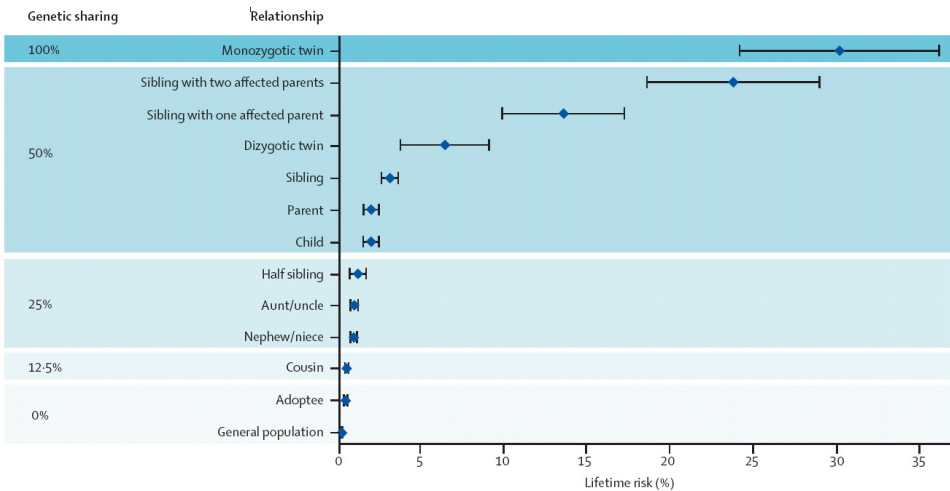


Figura 6. Riesgos de recurrencia para esclerosis múltiple en familias. Riesgos de recurrencia corregidos por edad para diferentes familiares de probandos con EM y grado del componente genético compartido entre el familiar y el probando. Extraída de Compston *et al.*⁹

6.2.2. Genes del complejo MHC

El complejo principal de histocompatibilidad (MHC, *Major Histocompatibility Complex*) es una región génica de 3,6 Mb que se localiza en el locus 6p21.3, en el brazo corto del cromosoma 6, y que se caracteriza por una alta densidad génica; una baja tasa de recombinación, lo que implica un elevado LD entre segmentos del MHC; y un elevado polimorfismo (o variabilidad genética)⁷⁷ (Figura 7). Esta región se divide en 3 subregiones: clase II, clase III y clase I; en función de su localización, de la más centromérica a la más telomérica, respectivamente. En ella existen alrededor de 421 genes, de los cuales se expresan el 60% y aproximadamente el 22% tienen funciones inmunorreguladoras.

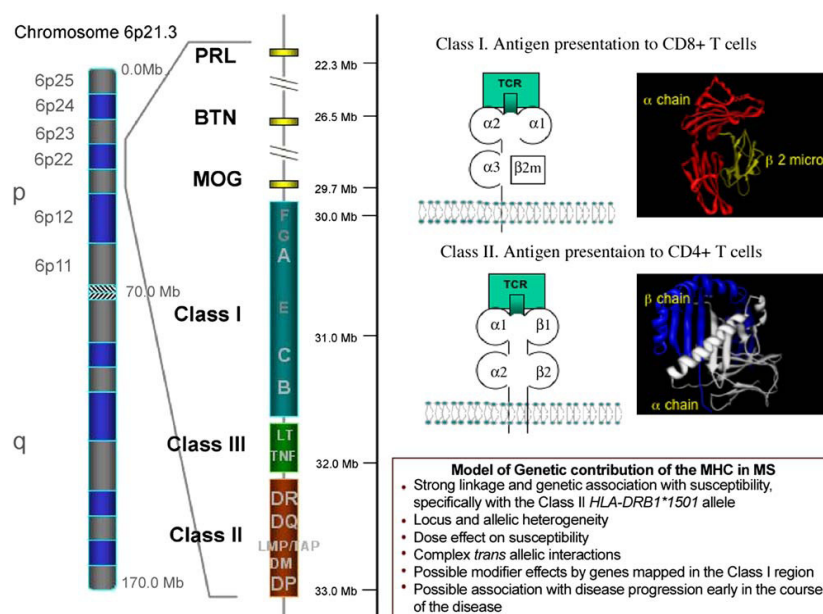


Figura 7. La región cromosómica 6p21.3 y EM. La secuencia completa de la región del MHC se completó y publicó en 1999. La subregión de clase I y la de clase II codifican para glicoproteínas de superficie celular altamente polimórficas, que intervienen en la presentación de antígeno a los LTs. La región del HLA clase III contiene muchos genes que codifican proteínas que no están relacionadas con la inmunidad celular, pero que modulan o regulan la respuesta inmune de otras formas. Figura extraída de Hauser *et al.*⁷⁸

En la década de los 70 se publicaron los primeros estudios que vinculaban los antígenos HLA con EM⁷⁹⁻⁸¹. En estos estudios se describió por primera vez la asociación de la EM con el alelo *HLA-DR2* del gen *HLA-DRB1*. Posteriormente, mediante técnicas de biología molecular se pudo concretar que el subtipo *HLA-DR2* que confiere susceptibilidad a EM es el *HLA-DRB1*15:01*, descrito hasta el momento como el principal factor genético asociado a EM⁸². Se han postulado varios mecanismos por los que el alelo *HLA-DRB1*15:01* puede estar implicado en la aparición de la EM. Uno de ellos es que este subtipo concreto presente algún péptido del SNC de manera muy efectiva a los linfocitos autorreactivos que no fueron eliminados en la selección tímica y que, de esta manera, se induzca una respuesta inflamatoria que lleve a la desmielinización de los axones neuronales^{83, 84}.

6.2.3. Genes fuera del MHC: aproximaciones genéticas y sus limitaciones

La presencia de alelos HLA ligados a la susceptibilidad a padecer EM no es suficiente para explicar todo el componente genético relacionado con la enfermedad. Esto pone en evidencia que la genética de la EM es compleja y que deben existir otros factores genéticos que contribuyen a desarrollar la enfermedad. Hay diferentes razones por las que hasta muy recientemente no se han detectado otros genes de riesgo fuera de la región HLA. Por un lado, el insuficiente tamaño muestral hacía imposible alcanzar una potencia estadística suficiente para detectar alelos con un efecto pequeño. Por otro lado, las poblaciones estudiadas en todo el mundo son bastante heterogéneas⁸⁵. A continuación, se mencionan los distintos abordajes realizados para identificar genes asociados con la EM fuera de la región HLA.

6.2.3.1. Estudios de ligamiento

Los estudios de ligamiento buscan la cosegregación de una región genómica en familias *multiplex*, es decir, familias con varios afectados de una enfermedad, en dos o más generaciones; y han sido particularmente exitosos en encontrar genes relacionados con desórdenes monogénicos o de transmisión mendeliana. En la EM, un 20% de los enfermos presenta un pariente con la enfermedad, pero las familias con más de cuatro miembros afectados en más de una generación son extremadamente raras⁸⁶, lo que limita el uso de los estudios de ligamiento en familias *multiplex* en EM. Los primeros estudios de ligamiento intentaron combinar señales de ligamiento en muchas familias pequeñas (principalmente parejas de hermanos afectados) empleando mapas de ligamiento basados en microsatélites (repeticiones de 2-10 nucleótidos), marcadores altamente polimórficos distribuidos a lo largo del genoma.

En la década de los 90, análisis de ligamiento mostraron que la EM estaba asociada con la región HLA^{82, 87, 88}. Todos estos estudios de ligamento en EM y otros posteriores fallaron en identificar regiones de asociación cromosómica fuera de esta región, a pesar de que uno de los estudios incluía 730 familias *multiplex* con 2.692 individuos⁸⁹. La falta de éxito está relacionada con la ya mencionada naturaleza compleja de la enfermedad en la que ningún gen es necesario ni suficiente para padecerla y la ausencia de grandes familias con

múltiples afectados. Sin embargo, la información obtenida a través de este tipo de estudios alentaba a abordar otras estrategias para identificarlos ⁹⁰.

6.2.3.2. Estudios de asociación

Los estudios de asociación se basan en la identificación de variantes que se asocian con la enfermedad ya sea por su implicación en la enfermedad o porque se encuentran en desequilibrio de ligamiento (LD, *linkage disequilibrium*) con la variante causal. Para ello, se comparan las frecuencias alélicas y/o genotípicas de los marcadores utilizados entre enfermos con EM (casos) y sujetos sanos no emparentados genéticamente (controles) para detectar diferencias significativas entre los dos grupos e inferir una asociación con la enfermedad.

6.2.3.2.1. Estudios de asociación de genes candidatos

Los estudios de asociación de genes candidatos investigaron genes cuyas funciones biológicas sugerían que podrían desempeñar un papel en la fisiopatología de la EM. En otros casos, los genes seleccionados fueron genes asociados al riesgo a padecer otras enfermedades de base inmunológica, en base a que a la hipótesis de que ciertas regiones cromosómicas conferían predisposición a múltiples trastornos inmunomediados, apoyando la idea de que un grupo de genes contribuye a un espectro genético común ⁹¹⁻⁹³. Se demostró que los estudios de genes candidatos tenían una capacidad mayor para detectar alelos comunes con un efecto modesto en la susceptibilidad a la enfermedad que los estudios de ligamiento ⁹⁴, pero se basan en la elección de genes candidatos adecuados, y por lo tanto, siempre son dependientes de una hipótesis previa. Sin embargo, a pesar de décadas de esfuerzo, poco progreso se realizó en la identificación de alelos de riesgo relevantes fuera de la región HLA mediante este abordaje.

Un ejemplo del estudio de genes candidatos en EM es el gen de la apolipoproteína E (APOE), que es un gen pleomórfico con 3 alelos principales: *APOE2*, *APOE3* y *APOE4*. El gen *APOE* es el locus de riesgo más importante para enfermedad de Alzheimer y su estudio en EM ha estado motivado por publicaciones que identificaban la región 19q13, que contiene el gen *APOE*, como ligado a EM ⁹⁵; unido a su rol funcional en el transporte de lípidos, la

immunorregulación, la plasticidad neuronal y los mecanismos de reparación ⁹⁶. Sin embargo, los estudios de asociación del gen *APOE* en EM han producido resultados contradictorios hasta la fecha, con algunos estudios describiendo una asociación significativa mientras muchos otros eran incapaces de confirmar estas asociaciones. Por este motivo, ha sido necesario un estudio más concluyente de la asociación potencial entre los dos SNPs de la *APOE* estudiados más comúnmente y la susceptibilidad a EM en poblaciones de ascendencia europea.

El gen *CNR1* (*cannabinoid receptor 1*) constituye otro ejemplo de genes candidatos estudiados en EM. Está localizado en la región 6q14-q15 y codifica para el receptor de cannabinoides CB₁ ⁹⁷. Su estudio se plasmó en distintas publicaciones que señalaban una relación del sistema endocannabinoide y esta enfermedad y en estudios que se habían realizado en los modelos experimentales de EM. Por ejemplo, mediante un estudio de microsatélites con 534 sujetos en población española se mostró una asociación del gen *CNR1* con el riesgo a sufrir EMPP. Sin embargo, no se obtuvieron resultados concluyentes en EMRR ⁹⁸. Debido al papel del receptor CB₁ en funciones importantes en el SNC y a la potencial relevancia de los cannabinoides en la EM ⁹⁹, consideramos importante aclarar el papel de gen *CNR1* en EM utilizando una cohorte más grande de sujetos.

Finalmente, como se ha mencionado anteriormente, existen variantes de riesgo comunes a múltiples enfermedades mediadas por el sistema inmunitario. En este sentido, por ejemplo, destaca el gen *IL12B*, con un papel probado en psoriasis por un estudio de asociación de genes candidatos ¹⁰⁰ y replicado en poblaciones caucásicas y asiáticas ^{101, 102}. Además, un cribaje de SNPs no sinónimos mostró que este gen era un locus de riesgo común para la enfermedad de Crohn y la colitis ulcerosa ¹⁰³ y más tarde, se describió su asociación con espondilitis anquilosante ¹⁰⁴. En la EM no se conoce la implicación de este loci, pero podría tener un papel en la etiología de la enfermedad como ya ocurre en otras enfermedades complejas con un carácter autoinmune.

6.2.3.2.2. Estudios de asociación mediante barrido genómico (GWAS)

Los estudios de asociación mediante barrido genómico (GWAS, *genome wide association studies*) pretenden analizar simultáneamente todo el genoma en busca de loci asociados a la enfermedad. El “Proyecto HapMap”¹⁰⁵ definió alrededor de 11 millones de SNPs en el genoma humano y encontró una elevada correlación entre algunos de ellos, definiendo patrones de LD dentro del genoma. Los datos del HapMap permitieron el desarrollo del mapeo de alta densidad del genoma mediante la selección de un grupo de polimorfismos, conocidos como “*tag SNPs*”, utilizados como marcadores genéticos de diversidad, que proporcionan la máxima información sobre los SNPs no genotipados debido al LD, de manera que se puede reducir la cantidad de polimorfismos genotipados en un estudio sin disminuir sustancialmente la cantidad de información generada.

Los GWAS son estudios libres de hipótesis, es decir, son abordados sin ninguna hipótesis particular previa. En los últimos años, la tecnología basada en chips ha permitido el genotipado eficiente y simultáneo de varios cientos de miles de SNPs en los GWAS a través del genoma a un coste efectivo y ha permitido descubrir muchos loci de susceptibilidad con efectos pequeños sobre el riesgo no solo para la EM sino también para otras enfermedades complejas.

Sin embargo, existe un debate en curso acerca de las herramientas que se utilizan para analizar e interpretar la gran cantidad de datos generados por los GWAS. Por ejemplo, el problema de corregir para múltiples ensayos se tiene que abordar con mucho cuidado dado que se llevan a cabo múltiples ensayos simultáneamente. Así, los GWAS requieren unos estrictos umbrales de significación debido al gran número de SNPs genotipados (típicamente entre 500.000 y 1.000.000) necesarios para cubrir la mayor parte (>80%) de la variación común a lo largo de todo el genoma. El nivel de significación del genoma completo se ha fijado en $P < 5 \times 10^{-8}$, que representa la *corrección de Bonferroni* para un millón de ensayos¹⁰⁶. Se ha demostrado teórica y empíricamente que estos valores de P son altamente indicativos de asociaciones genuinas que se replican en estudios posteriores^{90, 107}. Todavía hay muchos más resultados de asociación con valores de P en la zona gris, entre $P < 0,05$ y este umbral del genoma completo, dejando abierta la cuestión sobre cómo interpretar estos resultados sugerentes de significación. Además, hasta ahora, solo se han analizado los SNPs

sin incluir interacciones gen-gen (salvo en ocasiones, entre genes asociados significativamente a nivel individual) o gen-ambiente ¹⁰⁸.

En total, se han realizado 14 GWAS hasta la actualidad para estudiar el componente genético de la EM (Tabla 2). El éxito de estos estudios se ha relacionado directamente con el número de individuos estudiados. Así, aquellos estudios con menos de 800 casos no fueron capaces de identificar nuevas asociaciones reproducibles, mientras que cada uno de los estudios con más individuos, añadía nuevos loci a la creciente lista de loci de susceptibilidad en EM. Sin embargo, se ha observado que los GWAS realizados en EM presentan cierta discordancia en sus resultados, en el sentido de que no siempre apuntan a los mismos polimorfismos (o polimorfismos equivalentes) asociados a la enfermedad. En algunos casos, esto se podría deber a la gran heterogeneidad clínica presente en la enfermedad. Por lo tanto, es de suma importancia replicar los resultados obtenidos en estos estudios e intentar llegar a entender si en el estudio inicial se describen falsos positivos o si realmente los polimorfismos se asocian a un grupo de pacientes con características clínicas y moleculares comunes de los cuales estaba enriquecida la población del estudio inicial. Además, los distintos GWAS se han realizado en diferentes poblaciones y el hecho de que exista heterogeneidad genética entre las distintas poblaciones puede hacer que un buen marcador de la variante etiológica en una población no lo sea en otra, por lo que se observa una asociación de un SNP a la EM en un GWAS determinado pero no en otro. Por este motivo, es interesante realizar estudios de meta-análisis de GWAS con los que se pueda analizar la heterogeneidad genética entre las poblaciones estudiadas ya que además, de esta forma, aumenta la potencia estadística y se pueden detectar efectos previamente no observados. Como prueba de la heterogeneidad genética entre las distintas poblaciones, podemos citar como ejemplo, el gen *CBLB* (*E3 ubiquitin-protein ligase*) localizado en la región 3q13.11 en el cromosoma 3, que no se asoció al riesgo a EM en ningún GWAS hasta el estudio realizado por Sanna *et al.* ¹⁰⁹ que incluía 2.657 casos y 2.877 controles de población de Cerdeña, donde se encontró que el SNP rs9657904, localizado en el intrón 1 de este gen, presentaba una asociación significativa con la enfermedad [$P = 1,60 \times 10^{-10}$; OR (IC 95%) = 1,40 (1,27-1,57)]. Posteriormente, esta asociación se confirmó en un estudio llevado a cabo con 2.901 sujetos de la Italia continental ¹¹⁰. Sin embargo, no se conocía el efecto de este gen en una población independiente con un *background* genético distinto.

Tabla 2. Estudios de asociación a nivel del genoma completo en esclerosis múltiple.

	Población	Casos/Controles	SNPs
IMSGC (2007) ¹¹¹	Reino Unido y E.E.U.U.	931/Padres †	334.923
WTCCC1 (2007) ¹¹²	Reino Unido	975/1.466	12.374
Comabella et al. (2008) ¹¹³	España	242/242	428.867
Aulchenko et al. (2008) ¹¹⁴	Países Bajos	45/195	250.000
Baranzini et al. (2009) ¹¹⁵	Varios	978/883	551.642
De Jager et al. (2009) ¹¹⁶	Estados Unidos	860/1720	709.690
ANZgene (2009) ¹¹⁷	Australia y Nueva Zelanda	1.618/3.413	302.098
Sanna et al. (2010) ¹⁰⁹	Cerdeña (Italia)	882/872	555.335
Nischwitz et al. (2010) ¹¹⁸	Alemania	590/825	300.000
Jakkula et al. (2010) ¹¹⁹	Finlandia	68/136	297.343
IMSGC y WTCCC2 (2011) ¹²⁰	Varios	9.772/1.7376	475.806
Patsopoulos et al. (2011) ¹²¹	Varios	1.453/2.176	906.600
Matesanz et al. (2012) ¹²²	España	296/801	130.903
Martinelli-Boneschi et al. (2012) ¹²³	Italia	197/234	277.866

La tabla muestra el número de casos y controles que fueron incluidos en la fase de cribado y el número de SNPs genotipados. Algunos estudios también estudiaron SNPs imputados en la fase de cribado y muchos estudios genotiparon SNPs seleccionados en una cohorte de replicación. IMSGC= *International Multiple Sclerosis Genetics Consortium* (Consortio Internacional sobre Genética de la Esclerosis Múltiple), WTCCC= *Wellcome Trust Case Control Consortium* (Consortio de Casos y Controles del *Wellcome Trust*), ANZgene= *Australia and New Zealand Multiple Sclerosis Genetics Consortium* (Consortio sobre Genética de la Esclerosis Múltiple de Australia y Nueva Zelanda). † Ambos progenitores.

En el 2011, como fruto de la colaboración entre el Consortio Internacional sobre Genética de la Esclerosis Múltiple (IMSGC, *International Multiple Sclerosis Genetics Consortium*) y el Consortio de Casos y Controles del *Wellcome Trust 2* (WTCCC2, *Wellcome Trust Case Control Consortium 2*), se publicó el GWAS con mayor número de individuos realizado hasta el momento ¹²⁰ (Figura 8). Aparte de algunas señales de asociación independientes en la región del HLA, este estudio informó de 57 loci que ejercen un riesgo bajo o moderado ¹²⁰. En este estudio, se confirmaron 23 asociaciones publicadas previamente y se identificaron 34 nuevas variantes asociadas: 29 con significación a nivel del genoma completo y cinco loci que fallaban en reunir dichos criterios de significación pero con un alta significación estadística ($P < 5 \times 10^{-7}$). Los SNPs asociados más

significativamente en esos cinco loci fueron rs228614, en el gen *MANBA* (*mannosidase, beta A, lysosomal*); rs630923, *upstream* del gen *CXCR5* (*chemokine C-X-C motif receptor 5*); rs2744148, *downstream* del gen *SOX8* (*sex determining region Y-box 8*); rs2744148, *downstream* del gen *RPS6KB1* (*ribosomal protein S6 kinase, 70 kDa, polypeptide 1*) y rs6062314, en el gen *ZBTB46* (*zinc finger and BTB domain containing 46*)¹²⁰. Se necesitaron esfuerzos de validación independiente para discernir el supuesto rol de estos loci en el riesgo a padecer EM, dada la carencia de significación a nivel de genoma completo, ya que en los recientes GWAS muchas más variantes de riesgo podrían quedar ocultas bajo los estrictos umbrales de significación del genoma completo.

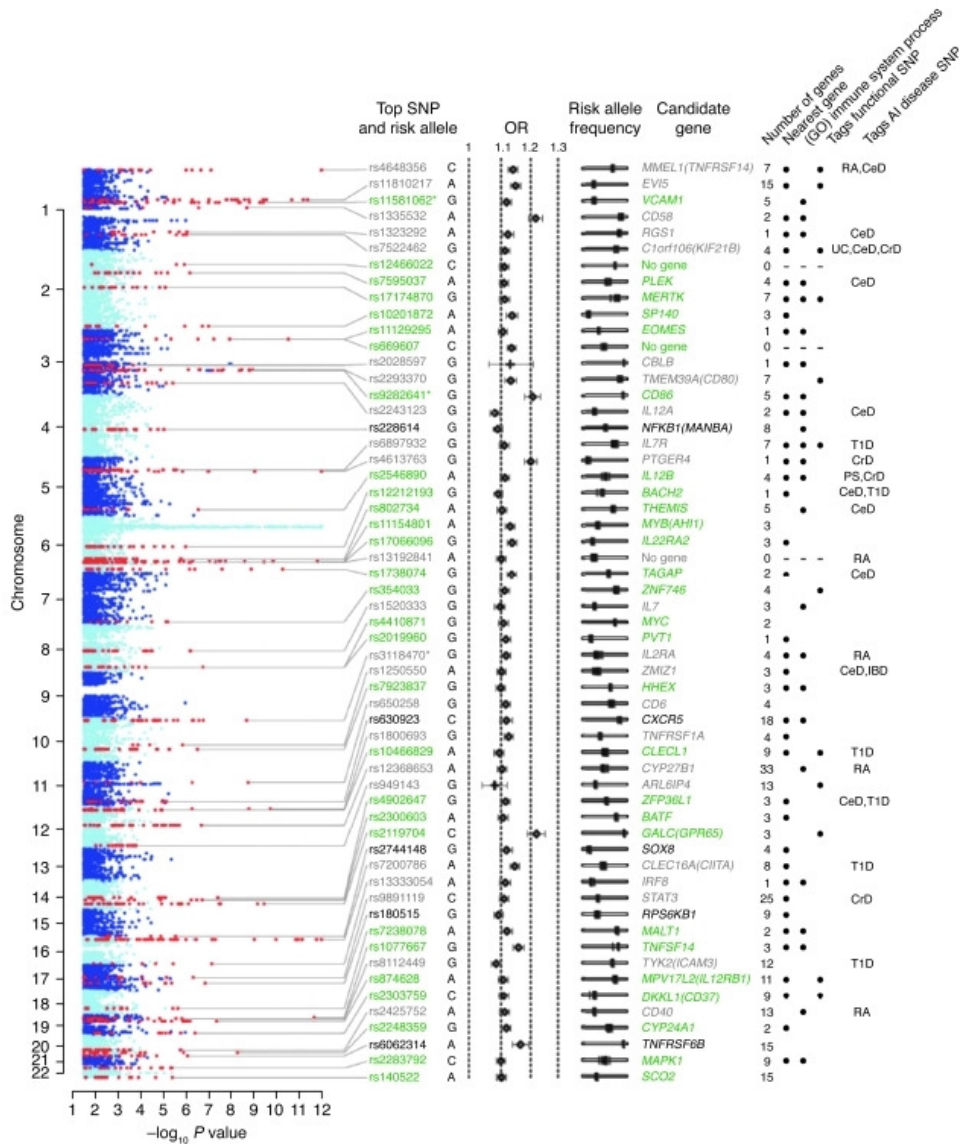


Figura 8. Regiones genómicas asociadas con riesgo a EM (Figura extraída de Sawcer *et al.* ¹²⁰). Esta lista de genes candidatos no-HLA relacionados con la EM se presentó después de un análisis refinado en el que se estudiaron 465.434 SNPs en 9.772 y 17.376 controles. Estos genes implicados ejercen efectos modestos a nivel poblacional con ORs comprendidas entre 1,1 y 1,3 ¹²⁰.

Respecto al GWAS realizado en el 2011 por Sawcer *et al.*¹²⁰ merece destacarse la coherencia en las características de estos 57 SNPs asociados, porque se ha observado que los genes que se localizan cerca de estos SNPs están significativamente sobrerrepresentados en términos relacionados con la inmunidad en el proyecto “Gene Ontology” (incluyendo la proliferación de linfocitos y la activación de células T)¹²⁰. Por otra parte, se ha demostrado que más de un tercio de estos SNPs (21 de 57) están asociados o están en LD con variantes asociadas con el riesgo a padecer otras enfermedades autoinmunes¹²⁰. Se ha observado que las distintas enfermedades autoinmunes comparten variantes genéticas de susceptibilidad, lo que sugiere un sustrato genético común, posiblemente debido a mecanismos moleculares comunes^{124, 125}. Los genes de susceptibilidad compartidos por muchas enfermedades autoinmunes pueden ser considerados como marcadores genéticos de autoinmunidad (Figura 9). Estos datos proporcionan una evidencia convincente del papel del sistema inmunitario en la etiología de la EM.

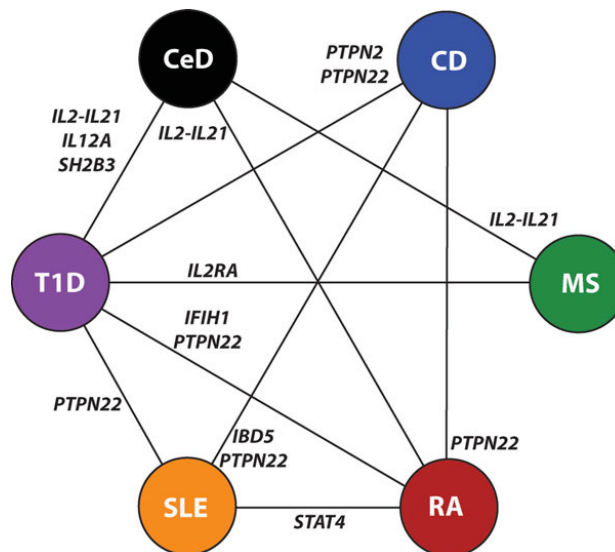


Figura 9. Genes de riesgo asociados a diversas enfermedades autoinmunes. Esta figura sugiere mecanismos comunes que conducen al desarrollo de enfermedades autoinmunes. No se incluyen los genes de la región HLA, pero variantes en esta región se asocian con todas las enfermedades representadas. CeD: enfermedad celiaca, CD: enfermedad de Crohn, T1D: diabetes mellitus tipo 1, MS: esclerosis múltiple, SLE: lupus eritematoso sistémico, RA: artritis reumatoide. Figura extraída de Lettre *et al.*¹²⁶

Por la información obtenida de los distintos GWAS se asume que la mayoría de las variantes genéticas que predisponen a la autoinmunidad influyen en la regulación inmunitaria y son responsables de la pérdida de los mecanismos que en condiciones normales se encargan del mantenimiento de la autotolerancia. Dichas variantes se suelen localizar en genes (o en regiones cromosómicas próximas a genes) que codifican citoquinas o sus receptores, receptores de patrones moleculares asociados a patógenos, factores de transcripción implicados en la diferenciación de células relacionadas con la respuesta inmunitaria o moléculas que participan o modulan la cascada de señalización de diversas citoquinas ¹²⁷. En este punto también hay que considerar que la EM tiene un componente neurodegenerativo y que se ha observado que la muerte oligodendrocitaria excitotóxica mediada por receptores de glutamato contribuye a las reacciones de desmielinización. Por estos motivos, al estudiar genes que codifican citoquinas, receptores de citoquinas, supresores de señalización de citoquinas, factores del complemento y receptores y transportadores de glutamato se podrían encontrar nuevos factores de riesgo a EM.

Finalmente, el fallo en descifrar el componente genético total de la susceptibilidad a EM también se puede deber en parte a la cobertura incompleta del genoma por los *arrays* disponibles utilizados en los GWAS. Prueba de ello, dos SNPs frecuentemente estudiados en la asociación con el riesgo a padecer EM en la era pre-GWAS, como son rs429358 (también conocido como “ε4”) y rs7412 (“ε2”) en el gen codificante de la APOE (Apolipoproteína E), están ausentes de muchas plataformas de genotipado de GWAS. Así, estos dos SNPs del gen *APOE* podrían no estar directamente estudiados en previos GWAS de EM ni por lo tanto, en los meta-análisis de los mismos. La más reciente liberación de datos de secuencia del genoma completo del “Proyecto 1.000 Genomas” hace ahora posible imputar genotipos de SNPs *proxys* de muchos GWAS. Esto podría ser importante para el estudio de la asociación entre los dos SNPs de la *APOE* estudiados más comúnmente y la susceptibilidad a EM en poblaciones de ascendencia europea.

7. Búsqueda de la “heredabilidad perdida” en la esclerosis múltiple

La heredabilidad es la proporción de la varianza fenotípica total que se debe a causas genéticas; en otras palabras, la heredabilidad mide la importancia relativa de la varianza genética como determinante de la varianza fenotípica. A pesar del gran número de variantes de riesgo identificadas en EM en los últimos años, la proporción de heredabilidad que estas explican es modesta, debido a sus efectos de pequeño tamaño. Esto mismo se ha observado en otras muchas enfermedades genéticamente complejas. Por lo tanto, algunos autores han acuñado el término “heredabilidad pérdida”¹²⁸. Las estimaciones recientes sugieren que los loci de riesgo identificados en EM explican aproximadamente solo un cuarto (~27%) de la heredabilidad total, la mayor parte de la cual se atribuye al locus MHC¹²⁹.

Los GWAS presentan limitaciones que hacen necesario desarrollar nuevas alternativas de estudio. Hasta ahora los marcadores asociados presentaban una frecuencia > 5%, una baja penetrancia y un efecto moderado (ORs < 1,3). Para alcanzar una potencia estadística suficiente y detectar marcadores con una baja frecuencia se necesita un número muy elevado de pacientes y controles, para lo que se reúnen individuos de diferentes poblaciones, pero eso dificulta la detección de variantes específicas de una población. Además, el LD en algunas de las regiones asociadas no hace posible determinar cuál es la verdadera variante causal y en ocasiones tampoco cuál es el gen implicado, especialmente cuando los marcadores asociados se localizan en regiones con una alta densidad génica. Por lo tanto, son necesarios estudios genéticos y funcionales adicionales en estas regiones.

Una de las alternativas que se proponen para superar algunas limitaciones actuales de los GWAS, es llevar a cabo GWAS que incluyan un mayor número de individuos, meta-análisis de todas las colecciones de datos de los GWAS independientes realizados y posteriores estudios de replicación de las señales asociadas más significativamente. Sin embargo, a no ser que no se detectaran previamente debido a razones técnicas (por ejemplo, una captura insuficiente por los *micro-arrays* de los GWAS), estos loci de riesgo genético adicionales serán de variantes frecuentes, con MAF (*minor allele frequency*) $\geq 5\%$ que ejercen efectos pequeños o muy pequeños (ORs < 1,10); o de variantes menos frecuentes con MAF < 5% y $\geq 0,5\%$ ejerciendo efectos ligeramente mayores ($1,2 < \text{ORs} < 2$) que actualmente podrían no superar el estricto umbral de significación establecido en los GWAS, como se ha

demostrado previamente ¹³⁰. En cualquier caso, los loci identificados mejorarán solo moderadamente la proporción de la heredabilidad explicada.

Los recientes GWAS, salvo algunas excepciones, no han analizado variantes raras con $MAF < 5\%$ en la población general, que podrían ejercer efectos mucho mayores con respecto a las variantes comunes y se piensa que una considerable fracción de la heredabilidad perdida en la EM podría explicarse por la presencia de dichas variantes ^{128, 131}. Además, los GWAS a pesar de su tamaño muestral, carecen probablemente de la potencia suficiente para alcanzar significación a nivel de genoma completo en variantes raras. Por otro lado, la mayoría de estas variantes no se conoce, a no ser que se hayan estudiado por secuenciación. Sin embargo, en desacuerdo con esta hipótesis, en un estudio de 25 loci asociados en los GWAS realizados en distintas enfermedades autoinmunes (incluyendo 10 loci de EM) con secuenciación de nueva generación, no se encontraron variantes raras de alto riesgo en ninguna de las regiones estudiadas ¹³². Además, un primer estudio en EM de una variante rara de alto riesgo en el gen *CYP27B1* que activa la vitamina D ¹³³, no fue confirmado posteriormente en estudios de validación independientes ^{134, 135}. Los futuros GWAS necesitarán estudiar un tamaño de muestras masivo para evaluar de forma concluyente el rol de las variantes raras en la EM.

Parte de la heredabilidad perdida puede deberse al efecto de variaciones genómicas estructurales. Estas variaciones son regiones de ADN que pueden abarcar desde una kb a algunas Mb e incluyen las inversiones y las translocaciones equilibradas; y los desequilibrios genómicos (inserciones o deleciones), comúnmente conocidas como variaciones de número de copia. Aunque tales variaciones han sido en gran parte inexploradas en relación a fenotipos complejos, existen estudios que señalaban la posible asociación de la EM con el gen *LILRA3* (*leukocyte immunoglobulin-like receptor A3 family*), caracterizado por la presencia de una delección de 6,7 kb que elimina los primeros 6 del total de 7 exones y que conduce a la pérdida de todos los dominios del tipo de las inmunoglobulinas del gen, lo que resulta en un alelo nulo ^{136, 137}. La asociación de la forma delecionada del gen *LILRA3* con la susceptibilidad a EM, se ha estudiado en varias ocasiones ¹³⁸⁻¹⁴¹, pero los resultados han sido poco concluyentes, con algunos estudios informando efectos significativos mientras otros son incapaces de confirmar dichas asociaciones. Por lo tanto, el papel del gen *LILRA3* continúa sin estar claro en la literatura.

Las modificaciones epigenéticas del ADN, entre las que se incluyen la metilación del ADN, la modificación de histonas y el silenciamiento de genes post-transcripcional asociado a los miRNAs ¹⁴², podrían ser razonables candidatos en la búsqueda de la “heredabilidad perdida”. Sin embargo, la evidencia a favor de efectos epigenéticos transgeneracionales en enfermedades genéticamente complejas está limitado actualmente a resultados en modelos animales y a muy pocos casos en estudios epidemiológicos ^{143, 144}.

Uno de los factores que podría estar involucrado en el porcentaje de “heredabilidad perdida” son las interacciones genéticas o epistasis. Este tipo de interacciones hacen referencia a la situación donde una única característica (fenotipo) se rige por dos o más genes y cada gen afecta la expresión de los otros genes involucrados. Marian *et al.* ¹⁴⁵ apunta a que una parte significativa de la “heredabilidad perdida” podría ser debida a una sobreestimación de la heredabilidad total consecuencia en parte de interacciones gen-gen más que debido a variantes aún no descubiertas. De ahí que Zuk *et al.* ¹⁴⁶ indiquen que se deben llevar a cabo estudios de epistasis para ver hasta qué nivel influyen en la heredabilidad de la enfermedad. Puede que en ocasiones la interacción no sea exclusivamente entre dos genes sino entre un grupo de genes, lo que hace más difícil su estudio ¹⁴⁷. Algunos estudios han examinado el papel de las interacciones epistáticas en la EM, particularmente las interacciones entre genes en la región HLA. Por ejemplo, en el estudio publicado por Ordonez *et al.* ¹⁴⁰ se describió que el gen *LILRA3* y el alelo *HLA-DRB1*15:01* actúan sinérgicamente para incrementar el riesgo a padecer EMRR.

Las interacciones entre factores genéticos y ambientales pueden ser otros factores interesantes para explicar al menos parte de la “heredabilidad perdida”: una variable genética particular en una situación ambiental adecuada puede ejercer un efecto significativamente mayor causando la enfermedad que si la exposición ambiental no estuviera presente. Hasta el momento, la atención se ha centrado en gran medida en el alelo *HLA-DRB1*15:01*. Por lo tanto, el efecto de una exposición ambiental sobre el desarrollo de la EM podría depender de si una persona presenta o no este alelo. Por otro lado, el riesgo causado por el alelo *HLA-DRB1*15:01* podría depender de ciertos factores ambientales. Por ejemplo, algunos estudios epidemiológicos en EM han descrito la interacción entre *HLA-DRB1*15:01* y factores ambientales como el virus de Epstein-Barr, el hecho de fumar y la vitamina D aunque su papel todavía no está claro ¹⁴⁸⁻¹⁵². Estos datos sugieren que las interacciones gen-ambiente desempeñan un papel importante en la EM y podrían explicar algo de la “heredabilidad

perdida”. Probablemente parte de ésta resida en interacciones gen-gen y gen-ambiente. Sin embargo, su estudio entraña gran dificultad debido a la multitud de factores que hay que tener en cuenta y a que estos pueden variar de unas poblaciones a otras. Sin embargo, dada la compleja estructura genética de la EM, si los efectos de estas interacciones no se incorporan en los análisis genéticos probablemente no emergerá una comprensión total de la enfermedad.

Los análisis basados en rutas funcionales también son considerados como una estrategia útil para explicar la “heredabilidad perdida”. Esta aproximación permite considerar las variantes genéticas asociadas con la enfermedad dentro del contexto de rutas biológicas conocidas y puede ofrecer un mayor conocimiento de los procesos que conducen a la enfermedad, porque si dichas variantes se agrupan dentro de una ruta de señalización conocida, las alteraciones de esta ruta podrían estar involucradas en el proceso que conduce a la enfermedad. Hasta la fecha han sido pocos los análisis de rutas publicados en EM. En un análisis integral de los datos de los GWAS, en el que se tuvieron en cuenta todos los SNPs con evidencia nominal de asociación ($p < 0,05$) ¹⁵³, se encontraron tres rutas que estaban sobrerrepresentadas en la EM: rutas de adhesión celular, rutas de comunicación y señalización celular; y una vía neuronal. Recientemente, en el GWAS del IMSGC/WTCCC2 mediante un análisis de la vía de diferenciación de los linfocitos Th, se encontró una sobrerrepresentación de los genes involucrados en la maduración de células T. Estos descubrimientos indican que las rutas de señalización de LTs pueden ser importantes en la patogénesis de la EM ¹²⁰.

Finalmente, hay otros aspectos a tener en cuenta cuando tratamos de explicar la “heredabilidad perdida”. Uno de ellos es la heterogeneidad clínica y paraclínica de la EM, que puede ser consecuencia, al menos en parte, de diferencias en la genética individual ¹⁵⁴. Sin embargo, los datos que apoyan la presencia de diferentes perfiles de riesgo genético en la EM es aún limitado. El fallo de identificar asociaciones convincentes en muchos análisis de subgrupos puede deberse parcialmente a limitaciones en la potencia en algunos estudios, a una mala clasificación en lo que respecta a variables clínicas y, como se ha postulado recientemente, al uso de sistemas de cuantificación subóptima de parámetros clínicos ¹⁵⁵. Por lo tanto, estudios de asociación a nivel del genoma completo considerando otros subgrupos de EM y otras variables clínicas y paraclínicas clasificadas más apropiadamente podrían

arrojar luz en cuanto a factores de riesgo potenciales para las distintas formas clínicas de EM.

Objetivos

En el presente trabajo de investigación se plantearon los siguientes objetivos:

1. Buscar genes de susceptibilidad a EM realizando un cribado de 368 polimorfismos localizados en genes de citoquinas, receptores de citoquinas, factores de transducción de la señal y receptores ionotrópicos y transportadores de glutamato; y validar las señales significativas en una cohorte independiente de replicación.
2. Realizar un estudio de asociación a gran escala de dos polimorfismos en el gen candidato *APOE* con datos procedentes de distintas fuentes.
3. Analizar la posible influencia en el riesgo a padecer EM de genes que se habían asociado en estudios previos.
 - a. Estudio de asociación de genes candidatos seleccionados porque confieren susceptibilidad a otras enfermedades inmunomediadas y que por su papel pudieran estar implicados en la EM: *IL12B* y *CNR1*.
 - b. Estudio de asociación del gen *CBLB* descrito como factor de riesgo en un GWAS incluyendo solamente población italiana.
4. Validar polimorfismos descritos como factores de riesgo a EM en estudios de GWAS y señales que quedaron al borde de la significación:
 - a. Validar loci identificados en el primer GWAS: *PKN2*, *GTF2B*, *EPHA4*, *TMEM39A*, *YTHDF3*, *ANKFN1* y *PTPRM*.
 - b. Estudiar loci que no alcanzaron el umbral de significación a nivel de GWAS: *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* y *ZBTB46*.
5. Identificar la variante o haplotipo causal de susceptibilidad a EM dentro del gen *CD6* y estudiar su papel funcional.

6. Buscar factores genéticos de susceptibilidad a EM no incluidos en los GWAS y que puedan contribuir a explicar la “heredabilidad perdida”: gen *LILRA3*.

Materiales y Métodos

1. Pacientes y controles

En este trabajo se incluyó información de un máximo de 849 enfermos de EM de ascendencia española, recopilados consecutivamente en la Unidad de Esclerosis Múltiple del Servicio de Neurología del Hospital Clínico San Carlos (HCSC) de Madrid y diagnosticados atendiendo a los criterios de Poser *et al.*¹² o de McDonald *et al.*¹³ La población de individuos control estuvo formada por un máximo de 832 individuos de origen caucásico, recopilados consecutivamente a lo largo del estudio y procedentes fundamentalmente de donantes de sangre y empleados del HCSC que no referían enfermedades de base inmunológica ni antecedentes familiares de las mismas. Estos individuos constituyen lo que denominaremos nuestra colección de muestras, que presenta las siguientes características:

- **Sexo:** la EM afecta principalmente a mujeres en una proporción de 2:1 con respecto a los hombres en la población mundial². La distribución de nuestra población de pacientes sigue la tendencia esperada con un 65% de mujeres afectadas por EM frente al 59% de mujeres presentes dentro del grupo control. Para aproximadamente el 2% de los pacientes y controles no se conoce este dato.
- **Edad:** la edad de debut media de la enfermedad en nuestra población de pacientes de EM es $30,0 \pm 8,8$ años, con un rango distribuido entre los 4 y los 63 años. La edad media de los controles en el momento de la toma de muestras es de $41,2 \pm 16,9$ años.
- **Formas clínicas:** el 85% de los pacientes se incluyeron dentro de la forma EMRR, el 8% fueron diagnosticados con la forma EMPP y el 6% presentaban la forma clínica EMSP.

Los individuos de nuestra colección que participaron en los distintos estudios aquí presentados fueron incluidos después de firmar el correspondiente consentimiento informado, aprobado previamente por el Comité de Ética de Investigación Clínica del HCSC.

2. Preparación de las muestras: aislamiento, cuantificación y estimación de la pureza del DNA

El aislamiento de DNA se realizó por el método de *salting out* ¹⁵⁶ a partir de los leucocitos procedentes de 10 ml de sangre periférica extraída de cada uno de los sujetos de nuestra cohorte estudiados. Posteriormente, se cuantificó la concentración de DNA genómico por espectrofotometría (*NanoDrop 1000, Thermo Scientific, Wilmington, DE, USA*) midiendo la absorbancia a una longitud de onda de 260 nm y considerando que cada unidad de densidad óptica se corresponde con 50 ng/μl de DNA. Se estimó también la pureza del DNA genómico mediante el *ratio* entre la absorbancia a 260 nm y la absorbancia a 280 nm y cuyo valores estuvieron comprendidos entre 1,7-1,8. Finalmente, a partir de los valores de concentración obtenidos, se prepararon diluciones de trabajo en agua destilada a una concentración de 10 ng/μl.

3. Genotipado de las muestras

El análisis de polimorfismos de un solo nucleótido (SNPs) se realizó mediante PCR a tiempo real utilizando sondas Taqman y empleando el equipo *ABI PRISM 7900HT (Applied Biosystems, Foster City, CA, USA)*. Una sonda Taqman® es un oligonucleótido complementario a la secuencia del DNA que contiene el polimorfismo que queremos analizar y que se encuentra marcada en su extremo 5' con un fluorocromo y en su extremo 3' presenta una molécula secuestradora de fluorescencia o *quencher*. En cada PCR se añaden dos sondas marcadas en el extremo 5' con un fluorocromo diferente, FAM (6-carboxifluoresceína) o VIC (4, 7, 2', 4', 5', 7'-hexaclaro-6-carboxifluoresceína), específicas para cada uno de los alelos del SNP que se pretende genotipar. Cuando dichas sondas están intactas la fluorescencia emitida por el fluorocromo es captada por el *quencher* y no es detectable. Al realizar una PCR, la polimerasa (*TaqGold, Applied Biosystems*) avanza amplificando en sentido 5' → 3' y al encontrarse una zona hibridada la degrada gracias a su actividad 5' → 3' exonucleasa, de modo que la fluorescencia es liberada de la capacidad secuestradora del *quencher* y es entonces detectada por el equipo. De este modo, dependiendo de la fluorescencia observada podremos saber el alelo o los alelos presentes en la muestra en esta región amplificada (Figura 10).

Las condiciones estándar de la PCR utilizada para el genotipaje de SNPs mediante sondas Taqman fueron las siguientes: 2 min a 50°C (activación de la enzima amperasa), 10 min a 95°C (activación de la enzima polimerasa e inactivación de la enzima amperasa) y 40 ciclos consistentes en 15 s a 95°C (desnaturalización) y 1 min a 60°C (hibridación y elongación).

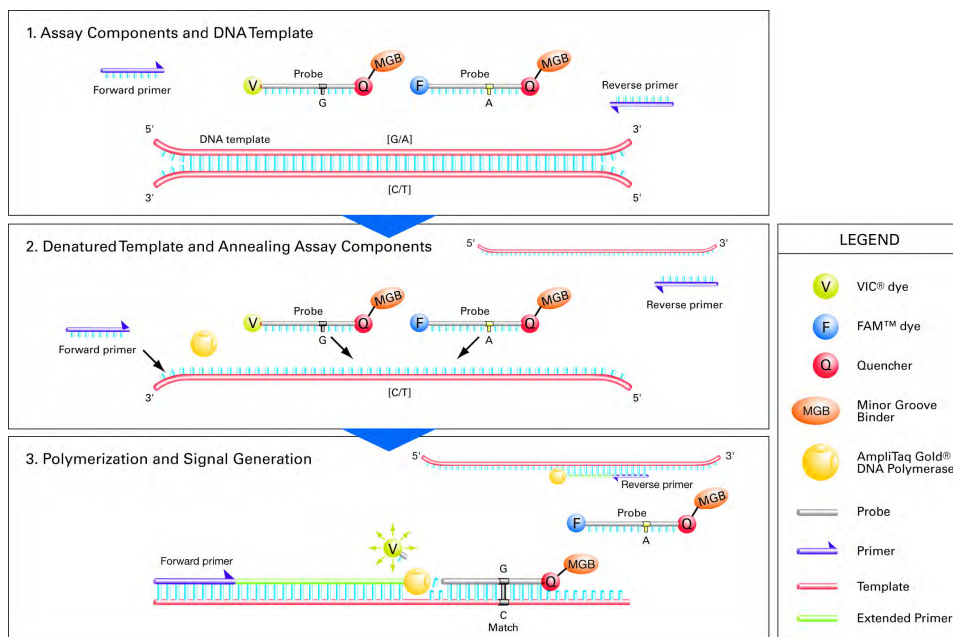


Figura 10. Esquema del proceso de la RT-PCR realizada con tecnología Taqman®. Cuando la sonda marcada encuentra el fragmento de DNA complementario, se unirá a él de manera específica, permitiendo que la polimerasa degrade la sonda y se produzca la liberación del *quencher*, emitiéndose fluorescencia. Imagen extraída de: https://tools.lifetechnologies.com/content/sfs/manuals/TaqMan_SNP_Genotyping_Assays_man.pdf

El subtipo *HLA-DRB1*15:01* del alelo *DR2* es el principal factor genético de susceptibilidad a EM. Para conocer si los individuos de nuestro estudio eran portadores de este alelo se realizó el genotipado del SNP rs3135388 también mediante tecnología Taqman®. Se ha descrito una correlación entre este SNP, situado en la región 3' del gen *HLA-DRA* y *HLA-DRB1*15:01* mayor del $r^2 = 0,94$ ¹⁵⁷ que fue confirmada en estudios

posteriores¹⁵⁸, ratificando a este SNP como buen marcador de la presencia/ausencia del alelo *HLA-DRB1*15:01*. En nuestra muestra se observó una concordancia del 97% entre el genotipado del SNP por tecnología Taqman® y la técnica PCR-SSOP (*Polymerase Chain Reaction-Sequence Specific Oligonucleotide Probe*).

El 34,6% de nuestros pacientes de EM son portadores del alelo *HLA-DRB1*15:01* frente al 14,8% en los controles, lo que confirma a este alelo como factor de susceptibilidad también en nuestra población.

4. Diseño de los estudios multicéntricos

El presente trabajo se ha llevado a cabo en colaboración con distintos centros de investigación nacionales e internacionales. En los siguientes apartados se describirán los detalles sobre el número de individuos incluidos en los distintos estudios así como las distintas estrategias de genotipado y los controles de calidad que se utilizaron.

4.1. Pacientes de esclerosis múltiple y controles

El diagnóstico de los pacientes con EM se estableció de acuerdo con los criterios diagnósticos estándar^{12, 13}. Todos los pacientes y los controles firmaron un consentimiento informado por escrito antes de ser incluidos en el estudio y los estudios se realizaron bajo la aprobación por parte de los correspondientes Comités Éticos de Investigación Clínica de cada uno de los centros. Los detalles de las colecciones de muestras utilizadas se muestran en la Tabla 3.

Tabla 3. Número total de pacientes y controles genotipados en cada uno de los estudios.

Lill <i>et al.</i>¹⁵⁹				
	Casos	(% mujeres)	Controles	(% mujeres)
Alemania	2.973	70	2.973	60
España (Bilbao)	495	73	562	69
España (Granada, Sevilla)	1.125	60	1.068	68
España (Madrid)	804	n.d.	628	n.d.
Francia (París)	1.344	64	768	61
Total	6.741		5.999	
Varadé <i>et al.</i>¹⁶⁰				
	Casos	(% mujeres)	Controles	(% mujeres)
España (Barcelona)	460	n.d.	462	n.d.
España (Bilbao y San Sebastián)	777	n.d.	706	n.d.
España (Madrid)	535	n.d.	536	n.d.
España (Málaga, Granada y Sevilla)	1.092	n.d.	1.226	n.d.
Total	2.864		2.930	
Lill <i>et al.</i>¹⁶¹				
	Casos	(% mujeres)	Controles	(% mujeres)
Alemania	3.657	71	2.956	60
Australia	662	75	284	51
España (Madrid)	849	n.d.	702	n.d.
España (Barcelona)	523	64	860	52
España (Bilbao)	430	72	508	68
España (Málaga, Granada y Sevilla)	1.061	68	1.108	69
Francia	846	69	266	62
Países Bajos	471	72	2.081	56
Reino Unido*	958	75	1.916	50
Total	9.457		10.681	
Vandenbroeck <i>et al.</i>¹⁶²				
	Casos	(% mujeres)	Controles	(% mujeres)
España (Andalucía)	1.131	61	1.237	68
España (Barcelona)	708	63	825	53
España (Bilbao) †	462	71	470	71
España (Madrid)	680	64	832	53
Estados Unidos (UCSF, afroamericanos)	895	77	658	70
Estados Unidos (UCSF, caucásicos)	505	68	451	67
Total	4.381		4.473	

Ortiz <i>et al.</i> ¹⁶³				
	Casos	(% mujeres)	Controles	(% mujeres)
España (Andalucía)	233	72	433	81
España (Barcelona)	268	57	310	52
España (Madrid)	1.386	67	810	58
Total	1.887		1.553	
Swaminathan <i>et al.</i> ¹⁶⁴				
	Casos	(% mujeres)	Controles	(% mujeres)
España (Andalucía) †	1.119	63	1.340	63
España (Bilbao) #	573	74	565	69
España (San Sebastián) #	250	62	249	66
España (Madrid)	639	63	810	53
Estados Unidos (UCSF, caucásicos)	507	68	450	67
Total	3.088		3.414	

En cada estudio se incluye el número de individuos (casos y controles) de cada cohorte participante, así como el porcentaje de mujeres cuando se disponía de este dato. * La colección del Reino Unido constó de 958 tríos con un descendiente afectado que se analizaron con la prueba de desequilibrio de transmisión. † El cribado inicial se llevó a cabo en la colección de Bilbao, que comprendía 462 pacientes con EM residentes en el País Vasco (Hospital de Basurto) y 470 controles sanos proporcionados por el Biobanco Vasco de la Fundación Vasca de Innovación e Investigación Sanitaria. # El cribado inicial se llevó a cabo en la colección del País Vasco, que incluye muestras recogidas de dos centros: el Hospital de Basurto (Bilbao) y el Hospital de Donostia (San Sebastián). ‡ Incluye muestras recogidas de 5 centros: Hospital Virgen Macarena (Sevilla), Hospital Carlos Haya (Málaga), Hospital Clínico, Hospital Virgen de las Nieves y Banco de Sangre (Granada). n.d.: no disponible.

4.2. Búsqueda y abstracción de datos

Los estudios de asociación genética encaminados a investigar el papel del gen *APOE* en la susceptibilidad a EM en poblaciones de ascendencia europea se identificaron por una búsqueda en la base de datos PubMed del NCBI mediante combinaciones de la siguiente terminología “APOE o (*apolipoprotein E*) o AD2 (*Alzheimer Disease 2*) o LPG (*lipoprotein glomerulopathy*) o 19q13” y “multiple sclerosis”; y revisando referencias cruzadas en publicaciones relacionadas. Solo se consideraron los estudios de tipo caso-control en los que se investigaba la asociación de los polimorfismos rs429358 y rs7412 del *APOE* con el riesgo a EM, publicados en revistas en inglés y disponibles hasta el primero de julio del 2012.

Para el estudio de asociación del gen *LILRA3* con EM, la terminología utilizada fue “*LILRA3* o *ITL6 (immunoglobulin-like transcript 6)*” y “multiple sclerosis” y se consideraron solo aquellos estudios que habían sido publicados en inglés hasta junio del 2014.

Siempre que fue posible, se extrajeron los siguientes datos de cada publicación: el diseño del estudio, el origen geográfico y la etnia de los participantes; el número de casos y controles; la proporción de mujeres en casos y controles; la edad de aparición en casos; la edad de la toma de la muestra en casos y controles; los SNPs analizados; las distribuciones genotípicas o alélicas; la distribución del genotipo en consistencia con el equilibrio de Hardy-Weinberg (HWE) y los detalles en el protocolo de genotipado. Se excluyeron aquellos estudios caso-control que hacían referencia a datos de controles no genotipados en el mismo laboratorio que los casos. Finalmente, cuando se identificó un estudio con colecciones de datos solapantes, solo se incluyó uno de las mismos, preferiblemente aquel con un mayor tamaño muestral.

4.3. Análisis de los datos procedentes de los GWAS

Se solicitaron y obtuvieron genotipos crudos procedentes de 3 GWAS para el estudio de asociación del gen *APOE* y la susceptibilidad a EM en poblaciones de ascendencia europea, que siguiendo los filtros de control de calidad pre-imputación comprendieron los siguientes individuos:

- *International Multiple Sclerosis Genetic Consortium (IMSGC)*: 926 casos y 2.319 controles.
- *GeneMSA*: 986 casos y 902 controles (de las colecciones de Estados Unidos, Suiza y los Países Bajos).
- *ANZgene*: 1.618 casos y 1.988 controles.

4.3.1. Controles de calidad de los datos procedentes de los GWAS

Para el estudio de los 2 SNPs del gen *APOE*, se llevó a cabo una depuración de los datos con *PLINK* v.1.07 en <http://pngu.mgh.harvard.edu/~purcell/plink/>¹⁶⁵. Primero, se

depuraron los datos en base a la eficacia y calidad de genotipado excluyendo aquellos SNPs con una MAF menor de 0,01, una tasa de pérdidas mayor al 2% o desviaciones del HWE (con el test χ^2 de *Pearson* en *PLINK*) con valores de P menores a 1×10^{-6} . Además, se excluyeron las muestras individuales con una eficacia de genotipado menor del 95%. Después, se excluyeron las muestras duplicadas y los parientes de primer y segundo grado dentro de las colecciones de los GWAS. Tras los controles de calidad se obtuvieron datos relativos a 298.309, 507.883 y 293.416 SNPs del IMSGC, GeneMSA y ANZgene, respectivamente.

4.3.2. Imputación de los datos relativos a la APOE y análisis de asociación

La imputación de la región 19q13 (chr19: 44,900,009-45,999,823, hg19), que incluye los SNPs rs429358 y rs7412 pertenecientes al gen *APOE*, se basó en los paneles de datos ‘1000GP_hg19_Jun2011.PhaseI’ usando el programa *IMPUTE* v.2.0 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html).

4.4. Selección de SNPs

En este trabajo se realizó un estudio de 368 SNPs en 55 genes o grupos de genes para estudiar su asociación con EM en una población residente del País Vasco¹⁶². Los SNPs analizados se seleccionaron mediante la realización de un *tagging*. Esta estrategia se basa en el estudio del mínimo número de SNPs necesarios para recoger gran parte de la variabilidad de una región génica. Se basa en el LD entre los SNPs situados en esa región. El valor del coeficiente de correlación (r^2) mide el grado de LD entre polimorfismos, oscila entre 0 y 1, indicando un valor 1 la equivalencia entre los mismos y el valor igual a 0 nula correlación entre ellos. Al realizar un *tagging* se selecciona solo un SNP entre todos aquellos que presentan elevado r^2 entre ellos. Para la selección se empleo el algoritmo de múltiples marcadores implementado en HapMap en cohorte CEU (HapMap Release # 19). En este programa existe la posibilidad de elegir la mínima MAF de los SNPs a incluir en el *tagging*, incluir o excluir forzosamente determinados SNPs, así como excluir SNPs que sean equivalentes a un haplotipo formado por marcadores ya seleccionados (*aggressive tagging*). El número de SNPs seleccionados dependerá del umbral de r^2 y MAF que elijamos, en

nuestro estudio $r^2 \geq 0,8$ y $MAF \geq 0,2$. Además, se forzó la inclusión de 34 SNPs no sinónimos en la selección del *haptag*. En la Tabla 4, se muestran los genes o grupos de genes que se incluyeron en este análisis:

Tabla 4. Selección de SNPs en el estudio de cribado en genes o grupos de genes.

Cromosoma	Genes o grupos de genes	Nº SNPs
1	<i>IL22RA1-IL28RA</i>	9
	<i>GRIK3</i>	13
	<i>IL23R-IL12RB2</i>	22
	<i>IL19-IL20-IL24</i>	5
2	<i>IL1RL2-IL1RL1-IL18R1-IL18RAP</i>	16
	<i>IL1F7-IL1F9-IL1F6-IL1F8-IL1F5-IL1F10-IL1RN</i>	30
3	<i>IL17RE-IL17RC</i>	4
	<i>IL17RB</i>	4
	<i>IL12A</i>	3
4	<i>IL21</i>	3
	<i>IL15</i>	4
	<i>GRIA2</i>	3
5	<i>IL7R</i>	7
	<i>SLC1A3</i>	2
	<i>LIFR</i>	4
	<i>OSMR</i>	14
	<i>C9</i>	5
	<i>C6</i>	7
	<i>IL31RA-IL6ST</i>	13
	<i>IL13</i>	2
	<i>IL9</i>	2
	<i>IL17B</i>	2
	<i>GRIA1</i>	22
6	<i>IL17A-IL17F</i>	11
	<i>IL20RA</i>	10
	<i>IL22RA2</i>	4
8	<i>IL7</i>	2
9	<i>C5</i>	6
10	<i>IL15RA</i>	6
11	<i>SLC1A2</i>	14
	<i>IL18BP</i>	2

	GRIA4	11
	IL18	3
12	IL23A	2
	IL22	2
	SOCS2	4
	IL31	3
13	IL17D	5
14	IL25	3
15	IL16	8
16	IL32	4
	SOCS1	3
	IL21R	8
	IL17C	2
17	SOCS3	2
19	EBI3	2
	TMED1	1
	IL12RB1	3
	IL28B–IL28A–IL29	12
	IL11	4
21	GRIK1	13
22	IL17RA	6
	MIF	8
	LIF	3
	OSM	4

En total se incluyen 368 polimorfismos en genes de citoquinas, receptores de citoquinas, proteínas supresoras de la señalización por citoquinas, factores del complemento y receptores y transportadores del glutamato.

En el caso del estudio del gen *CD6* las premisas empleadas también fueron $r^2 \geq 0,8$ y $MAF \geq 0,2$. Se seleccionaron 13 SNPs marcadores de haplotipo para el estudio: 9 SNPs intrónicos, 3 SNPs no sinónimos incluidos forzosamente y el rs17824933 estudiado anteriormente. También se incluyó el SNP rs650258 localizado cerca del gen *CD6* que había mostrado una fuerte asociación con la susceptibilidad a EM ¹²⁰.

En otros casos, el estudio de los SNPs incluidos en un determinado estudio vino marcado por la información procedente de los GWAS en los que habían destacado como posibles factores de susceptibilidad a EM, por su asociación con otras enfermedades

autoinmunes o por estudios de genes candidatos realizados previamente. En la Tabla 5 se indican los SNPs que se han genotipado en los diferentes estudios y la estrategia de genotipado en cada caso. Para el estudio de 368 SNPs, que consta de un cribado inicial, la tabla solo incluye aquellos SNPs que superaron esta fase inicial y se replicaron en una cohorte independiente.

Tabla 5. Listado de los SNPs analizados en los diferentes estudios.

Lill et al.¹⁵⁹				
Región cromosómica	Gen	SNP	Localización (5'→3')	Cambio aminoacídico
19q13.32	<i>APOE</i>	rs429358 (C/T)	Exón 4	130 C/R
19q13.32	<i>APOE</i>	rs7412 (C/T)	Exón 4	176 R/C
Varadé et al.¹⁶⁰				
Región cromosómica	Gen	SNP	Localización (5'→3')	Cambio aminoacídico
1p22.2	<i>PKN2</i>	rs305217 (A/G)	Intrón 2	--
1p22.2	<i>GTF2B</i>	rs7538427 (C/T)	Intrón 2	--
2q35	<i>EPHA4</i>	rs1517440 (C/T)	--	--
3q13.33	<i>TMEM39A</i>	rs1132200 (C/T)	Exón 9	487 A/T
8q12.3	<i>YTHDF3</i>	rs12115114 (A/G)	--	--
17q22	<i>ANKFN1</i>	rs17758761 (A/C)	--	--
18p11.23	<i>PTPRM</i>	rs4798571 (A/G)	Intrón 1	--
3q13.11	<i>CBLB</i>	rs9657904 (C/T)	Intrón 1	--
5q33.3	<i>IL12B</i>	rs6887695 (C/G)	--	--
5q33.3	<i>IL12B</i>	rs10045431 (A/C)	--	--
6q15	<i>CNR1</i>	rs1049353 (C/T)	Exón 2	--
Lill et al.¹⁶¹				
Región cromosómica	Gen	SNP	Localización (5'→3')	Cambio aminoacídico
4q24	<i>MANBA</i>	rs228614 (A/G)	Intrón 12	--
11q23.3	<i>CXCR5</i>	rs630923 (A/C)	Promotor	--
16p13.3	<i>SOX8</i>	rs2744 (A/G)	--	--
17q23.1	<i>RPS6KB1</i>	rs180515 (A/G)	3' UTR	--
20q13.33	<i>ZBTB46</i>	rs6062314 (C/T)	Intrón 2	--

Vandenbroeck <i>et al.</i> ¹⁶²				
Región cromosómica	Gen	SNP	Localización (5'→3')	Cambio aminoacídico
16p13.13	<i>SOCS1</i>	rs243324 (A/G)	Promotor	--
1p36.11	<i>IL28RA</i>	rs1416834 (A/G)	Intrón 4	--
5p13.1	<i>OSMR</i>	rs3805558 (A/G)	Intrón 1	--
6q23.3	<i>IL22RA2</i>	rs202573 (A/G)	Intrón 4	--
16p13.13	<i>CLEC16A</i>	rs12708716 (A/G)	Intrón 18	--
16p13.13	<i>CLEC16A</i>	rs2903692 (A/G)	Intrón 21	--
16p13.14	<i>CLEC16A</i>	rs6498169 (A/G)	Intrón 21	--
Swaminathan <i>et al.</i> ¹⁶⁴				
Región cromosómica	Gen	SNP	Localización (5'→3')	Cambio aminoacídico
11q12.2	<i>CD6</i>	rs3019561 (G/T)	Intrón 1	--
11q12.2	<i>CD6</i>	rs3019562 (C/G)	Intrón 1	--
11q12.2	<i>CD6</i>	rs3019548 (C/T)	Intrón 1	--
11q12.2	<i>CD6</i>	rs2905506 (C/T)	Intrón 1	--
11q12.2	<i>CD6</i>	rs11230548 (A/C)	Intrón 1	--
11q12.2	<i>CD6</i>	rs17824933 (C/G)	Intrón 1	--
11q12.2	<i>CD6</i>	rs11230555 (A/C)	Intrón 1	--
11q12.2	<i>CD6</i>	rs916811 (A/G)	Intrón 1	--
11q12.2	<i>CD6</i>	rs11230559 (C/T)	Intrón 1	--
11q12.2	<i>CD6</i>	rs2237997 (C/T)	Intrón 1	--
11q12.2	<i>CD6</i>	rs11230563 (C/T)	Exón 4	225 R/W
11q12.2	<i>CD6</i>	rs2074225 (C/T)	Exón	257 A/V
11q12.2	<i>CD6</i>	rs12360861 (A/G)	Exón 5	271 A/T
11q12.2	<i>CD6</i>	rs650258 (C/T)	--	--

4.5. Genotipado de SNPs

Para el cribado de los 368 SNPs en genes de interés realizado en el estudio de Vandenbroeck *et al.*¹⁶², se diseñó un panel personalizado de SNPs en el Servicio de Atención al Cliente de *Illumina* (*San Diego, CA, USA*). Los SNPs se caracterizaron en cada muestra en placas de 384 siguiendo el protocolo *Golden Gate* de *Illumina*. Se usaron 250 ng (5 µl a 50 ng/µl) de DNA genómico por reacción de genotipado y las matrices se analizaron en el sistema *BeadStation-500GX* de *Illumina*. Los archivos de intensidad obtenidos se decodificaron en datos de genotipado con el software *BeadStudio v.3.0* (*Illumina*).

Para los otros casos, el genotipado de los diferentes SNPs se realizó en el propio centro participante o en centros especializados de genotipado en función de las necesidades particulares del estudio o de la decisión de los investigadores de cada localización. Uno de estos centros, es el **Centro Nacional de Genotipado** (CEGEN, en su nodo ubicado en Santiago de Compostela, <http://www.usc.es/cegen/>), donde el genotipado de los SNPs se realizó mediante la plataforma *iPLEXTM MassARRAY* (por tecnología *Sequenom*). Todas las muestras de DNA se enviaron en placas de 96 pocillos previa cuantificación usando *PicoGreen* (*Molecular Probes, Eugene, OR, USA*). Este proceso de genotipado consta de dos reacciones. En primer lugar se amplifican los fragmentos de DNA que contienen los SNPs de interés mediante una PCR multiplex y, a continuación, se lleva a cabo la reacción de discriminación alélica. En la tecnología *iPLEX[®] Gold* todas las reacciones terminan tras la extensión de una única base, que genera una diferencia de masa en el producto de extensión según la base añadida. Esta diferencia de masa se detecta mediante espectrofotometría de masas *MALDI-TOF*. Las reacciones se realizan en placas de 384 pocillos y los productos son transferidos de forma automatizada mediante un robot a la superficie del *Spectrochip*, que será leído en el espectrofotómetro de masas.

Otro centro especializado de genotipado al que también se enviaron muestras para ser analizadas fue el **Instituto de Genética Molecular Max Planck** en Berlín, donde se genotiparon las muestras alemanas usando secuenciación de Sanger. Los oligómeros de la PCR se diseñaron para abarcar en el mismo amplicón los dos polimorfismos del gen *APOE* incluidos en este estudio (*primer forward*: GCCTACAAATCGGAACTGGA, *primer reverse*: ACAGAATTTCGCCCCGGCCTGGTACTGCC). Después de la PCR, cada amplicón purificado se secuenció usando el kit comercial *BigDye terminator v3.1* (*Applied Biosystems*) y el resultado se analizó en un secuenciador *ABI3730XL* (*Applied Biosystems*). Los genotipos se determinaron con el software *Seqman II v.5.0* (*DNASTAR, Madison, WI, USA*).

A continuación se describen brevemente algunas de las estrategias de genotipado por las que se optó en otros de los centros participantes:

La plataforma *Open Array[®] Real-Time PCR* (*Applied BioSystems*) está basada en los ensayos de genotipado de las sondas Taqman[®] (tanto diseñadas como personalizadas) y la tecnología *Open Array[®]*. La tecnología *Open Array[®]* usa placas que contienen 48 *subarrays*, cada una con 64 pocillos.

El kit *Human610-Quad BeadChip* (Illumina) se basa en la tecnología *Infinium*®. Las dianas a estudio son capturadas y detectadas en *arrays beadchip* de alta densidad. Su formato de 4 muestras ofrece un incremento significativo del rendimiento de la muestra y reduce la manipulación en el laboratorio, ya que tiene 550.000 SNPs más unos 60.000 marcadores genéticos adicionales por muestra.

4.6. Controles de calidad

Debido al carácter multicéntrico de estos estudios, se realizaron diversas pruebas de control de calidad para garantizar que el genotipado de las muestras fuera concordante tanto dentro de un mismo centro como dentro de los distintos centros participantes. Entre ellas, se incluyen algunas de las siguientes:

- Se genotiparon muestras del CEPH (*Centre d'Etude du Polymorphisme Humain*) para descartar inconsistencias mendelianas que pudieran revelar el fallo de la técnica para cada SNP.
- Las muestras carentes de genotipos para más de dos SNPs en los diferentes estudios se excluyeron antes de realizar los análisis.
- Se intercambiaron muestras entre los distintos centros participantes para asegurar la concordancia en el genotipado. Las muestras se genotiparon sin el conocimiento previo del genotipo que presentaban en su localización de origen y los resultados fueron intercambiados posteriormente.
- El genotipado y la asignación de los genotipos se llevó a cabo de forma ciega al estatus fenotípico de los sujetos.
- Un 5-10% de las muestras, dependiendo del estudio, se genotiparon por duplicado de forma aleatoria. Los resultados fueron concordantes con una precisión media de más del 99,9%.

- Se analizaron SNPs con datos de alta calidad, con una eficacia de genotipado mayor al 95% salvo para el rs1049353 del gen *CNR1* (93% en controles), y para el rs916811 y el rs11230563 del gen *CD6* en la colección de Madrid, con un 94% y un 94,5%, respectivamente.
- Las desviaciones del HWE en controles ($P \leq 0,05$) se estudiaron mediante el test χ^2 de *Pearson*. Se analizaron aquellos SNPs cuyas frecuencias alélicas en controles y en los padres no afectados de los núcleos familiares se ajusta a los valores esperados bajo el HWE ($P > 0,05$).

5. Estrategia de genotipado del gen *LILRA3*

Las formas completa y delecionada del gen *LILRA3* se detectaron en un único tubo de PCR con tres *primers*: dos oligonucleótidos *forward* (FP1, FILT6i7+967, 5'-gacttgtaagggttaaaaagccaa-3' y FP2, FILT6e2-3638, 5'-catctcgatctgccactgacac-3') y uno *reverse* (RP, RILT6i7+1071, 5'-gacagcagattctaaaacagtgg-3'). Cien nanogramos de DNA se amplificaron con 15 picomoles de cada primer en 15 μ l de tampón de PCR (67 mM Tris-HCl, pH 8.8, 16 mM $(\text{NH}_4)_2\text{SO}_4$, 2 mM MgCl_2 , 0,01% Tween-20) con 0,3 U de TaqPolimerasa (*EcoTaq*, *Ecogen*, Madrid, España) y 100 mM de deoxirribonucleótidos trifosfato. Las condiciones de PCR fueron 95°C por 2 minutos; 10 ciclos de 10 segundos a 94°C y 40 segundos a 65°C; y 20 ciclos de 10 segundos a 94°C, 20 segundos a 61°C y 30 segundos a 72°C. En la forma completa de *LILRA3*, FP1 y RP producen un amplicón de 150 pb, mientras que FP2 hibrida con el DNA, pero no conduce a ningún producto debido a una longitud del amplicón excesiva (~7 kb). En la forma delecionada, FP2 y RP conducen a un producto de 241pb, mientras que se pierde el sitio diana de FP1. Los dos amplicones se separaron en geles de agarosa al 2% y se visualizaron con bromuro de etidio.

6. Muestras para separación de células activadas por fluorescencia (FACS) y ELISA

Para este estudio se utilizaron 10 ml de sangre heparinizada en litio extraída por venopunción de 27 pacientes con EM (Departamento de Neurología, Instituto de

Neurociencias Clínicas, Hospital Regional Universitario Carlos Haya, Málaga) y de 12 controles sanos (University Klinikum Eppendorf, Hamburgo, Alemania). Las características clínicas de estos 27 pacientes con EM están reseñadas en la Tabla 6. Las células mononucleares de sangre periférica (PBMC) se purificaron usando un gradiente Ficoll-Hypaque, como se describía en el protocolo del proveedor (*ICN Biomedicals Inc., OH, USA*), se contaron e inmediatamente se utilizaron o se criopreservaron en la presencia de dimetil sulfóxido (DMSO) 10% (v/v), y 20% de suero de ternera fetal, para posteriormente almacenarse a -196°C en nitrógeno líquido hasta su utilización. Cuando se utilizaron PBMCs congeladas, éstas se descongelaron inmediatamente antes de su uso.

Tabla 6. Variables clínicas y demográficas de los pacientes con EM de los que se recogieron PBMCs para el estudio funcional.

Características	Casos	Controles
Mujer/Hombre (% Mujeres)	18/9 (66,6)	7/5 (58,3)
Media de edad en años ± ES (Rango)	43,1 ± 11,9 (23-66)	33,08 ± 6,68 (25-45)
Pacientes con EMRR (%)	23 (85,2)	-
Pacientes con EMSP (%)	4 (14,8)	-
Media de edad al inicio de la enfermedad en años ± ES (Rango)	27,9 ± 10,4 (15-55)	-
Duración media de la enfermedad en años ± ES (Rango)	15,1 ± 6,6 (5-28)	-
EDSS actual ± ES (Rango)	2,22 ± 1,21 (0-5)	-

7. Citometría de flujo para estudio de expresión en superficie celular

La lista de anticuerpos monoclonales y la combinación de marcadores usados para identificar los diferentes grupos celulares y sus proveedores se encuentran en la Tabla 7 y en la Tabla 8. En el estudio se incluyó un tubo sin marcaje y los correspondientes controles FMO. CD3 y CD56 se usaron para identificar subgrupos de linfocitos: LTs ($\text{CD3}^+ \text{CD56}^-$), NK ($\text{CD3}^- \text{CD56}^+$) y NKT ($\text{CD3}^+ \text{CD56}^+$). Dentro de los LTs, CD4 y CD8 marcan células cooperadoras y citotóxicas, respectivamente. Posteriormente, CD45RA, CD27 y CD28 se usaron para distinguir células vírgenes ($\text{CD45RA}^+ \text{CD27}^+ \text{CD28}^+$), células de memoria centrales ($\text{CD45RA}^- \text{CD27}^+$) y células de memoria efectoras ($\text{CD45RA}^- \text{CD27}^- \text{CD28}^+$), así

como las células de memoria efectoras diferenciadas terminalmente (TEMRA, CD45RA⁺ CD27⁻ CD28⁻)^{166, 167}. Las células NK se pueden dividir adicionalmente en las clásicas células NK citotóxicas CD56^{dim} (CD56^{int} CD16⁺), y células NK inmunorreguladoras CD56^{bright} (CD56^{hi} CD16⁺)¹⁶⁸. Para el marcaje, las PBMCs se descongelaron, se lavaron 2 veces y se incubaron durante 45 minutos con los cócteles de anticuerpos señalados en la Tabla 8. Posteriormente, las células se lavaron, resuspendieron en 300 µl de tampón FACS y se analizaron en el citómetro FACS Canto (BD Bioscience, San Diego, CA). El software FACS Diva v.6.1.3 (BD Bioscience) se utilizó para analizar los datos crudos.

Tabla 7. Detalles de los anticuerpos monoclonales usados en citometría de flujo.

Especificidad	Marcador fluorescente	Clon	Tipo	Casa comercial
CD3	PeCy5.5	UCHT1	primario	Biozol
CD8a	PacBlue	RPA-T8	primario	Biologend
CD4	PECy7	RPA-T4	primario	BD Biosciences
CD4	AF488	RPA-T4	primario	BD Biosciences
CD56	APC	N901 (NKH-1)	primario	Beckman Coulter
CD16	APCCy7	3G8	primario	Biologend
CD28	PECy7	CD28.2	primario	Biologend
CD27	APCCy7	M-T271	primario	BD Biosciences
CD45RA	APC	HI100	primario	eBiosciences
CD6 (D1)	PE	M-T605	primario	BD Biosciences
161.8	Sin marcaje		*	

* Proporcionado por el Prof. Francisco Lozano, IDIBAPS, Facultad de Medicina, Universidad de Barcelona, Barcelona.

Tabla 8. Combinación de anticuerpos monoclonales usados para el marcaje de antígenos de la superficie celular.

Subgrupo celular	Marcador fluorescente					
	V450	PE	PE-Cy5.5	PE-Cy7	APC	APC-Cy7
LT-Efector	CD8	CD6	CD3	CD28	CD45RA	CD27
NK	CD8	CD6	CD3	CD4	CD56	CD16

8. Ensayos de cultivo celular, proliferación y ELISA para la cuantificación de citoquinas

Para estos ensayos, se resuspendieron 10^6 PBMCs de cada donante en 500 μ l de medio RPMI (GIBCO, Life technologies, Grand Island, NY, USA) con 5% de suero fetal bovino (Biochrom AG, Berlín, Germany), se alicuotaron en una placa de 48 pocillos (Greiner Bio-One, Stonehouse, UK) y se cultivaron bajo 3 condiciones diferentes: sin estímulo y estimulado con anti-CD3 (OKT3, 100 ng/ml) con o sin anticuerpo anti-CD6 (161.8, concentración 10 μ g/ml). Luego, las PBMCs de los controles sanos se marcaron con 2 μ M eFluor670 (eBioscience, San Diego, CA, USA) de acuerdo con el protocolo del fabricante para determinar la proliferación celular. De forma resumida, el eFluor de reconstitución y las diluciones se realizaron en tampón salino a 4°C y la incubación con el marcador se llevó a cabo a temperatura ambiente durante 10 minutos. Después, se añadió medio estándar para detener el marcaje y se incubó en hielo durante 5 minutos. Las células marcadas se transfirieron a placas de 48 pocillos con el respectivo estímulo para los cultivos celulares.

Para estudiar por citometría de flujo el porcentaje de células que habían sufrido división a día 0 y día 3, estas se incubaron con marcadores específicos de linaje. Los sobrenadantes de los cultivos se recogieron al tercer día y se cuantificaron los niveles de citoquinas IL-17 e IFN- γ por ELISA usando los kits *Ready-Set-Go* (eBiosciences). La concentración de citoquinas (pg/ml) se calculó usando las curvas patrón generadas usando los respectivos calibradores.

9. Análisis de datos

9.1. Potencia y corrección múltiple

La potencia estadística de un estudio se define como la capacidad de detectar de forma significativa una asociación real, en nuestro caso, entre un alelo y el fenotipo EM. Depende de diversos factores: el tamaño de la muestra, el efecto que el alelo asociado tiene sobre la población (OR) y la MAF; de forma que al aumentar los dos primeros o la frecuencia alélica acercarse a valores intermedios, aumentará la potencia estadística de nuestro estudio. La potencia se estimó usando el *Genetic Power Calculator*

(<http://pngu.mgh.harvard.edu/~purcell/gpc/cc2.html>)¹⁶⁹ o la calculadora de potencia de CATS (<http://www.sph.umich.edu/csg/abecasis/CaTS>)¹⁷⁰, en base a las frecuencias alélicas de los SNPs estudiados para poblaciones control de ascendencia europea.

Un suceso se considera estadísticamente significativo cuando la probabilidad de que aparezca por azar es menor del 5%, sin embargo, cuando realizamos un alto número de comparaciones es inevitable que aumente el número de sucesos significativos por azar, lo cual incrementa la posibilidad de tener falsos positivos. Para evitar este suceso es necesario hacer correcciones como la clásicamente utilizada *corrección de Bonferroni*. Esta corrección sostiene que para mantener la probabilidad global ($P < 0,05$) de rechazar incorrectamente la hipótesis nula, el nivel de corte a utilizar debe ser de $0,05/K$ si se efectúan un número K de comparaciones independientes. En nuestro estudio, hemos empleado esta corrección en SNPs independientes genotipados en un mismo estudio. Sin embargo, esta corrección resulta muy extrema en el caso de polimorfismos que estén en una misma región de LD (ya que no se puede definir la independencia de los mismos), así como en el caso de replicaciones de SNPs previamente asociadas a EM, en cuyo caso esta corrección se hace innecesaria. En el caso de los estudios GWAS, el uso de correcciones es fundamental y se constituye en función del número de SNPs analizados. Habitualmente, se establece como corrección adecuada un nivel de corte inferior a 10^{-6} - 10^{-7} ¹⁷¹.

9.2. Estudios caso-control

Mediante estos estudios se comparan la frecuencia de un factor considerado de riesgo en un grupo de enfermos (en nuestro caso EM) con la existente en otro grupo de individuos libres de enfermedad (grupo control). La comparación de las frecuencias genotípicas, alélicas y haplotípicas entre enfermos y controles o entre los diferentes subgrupos de pacientes se realizó mediante el test χ^2 o el test exacto de Fisher (cuando los valores esperados fueron menores de 5), utilizando el programa *Statcalc* (*EpiInfo* v.6). Además, en las comparaciones 2x2, la intensidad de la asociación o el efecto se calculó mediante la *odds ratio* (OR) y su intervalo de confianza (IC) se calculó mediante el método de *Woolf*.

9.3. Estudio de haplotipos

El LD se produce cuando dos o más combinaciones de alelos se heredan juntas con una frecuencia superior a la esperada por azar. La combinación de estos alelos constituye los denominados haplotipos. Para el estudio de haplotipos, en el estudio de mapeo y análisis funcional del gen *CD6*, empleamos el algoritmo de Expectación-Maximización implementado en el programa *Haploview*, el cual estima los diferentes haplotipos presentes en una población (así como sus frecuencias) a partir de las frecuencias genotípicas de los SNPs en la muestra. La comparación posterior entre las frecuencias haplotípicas en enfermos y controles se realiza con el test χ^2 ¹⁷².

Otro de los análisis que se realizó para el estudio de haplotipos fue el análisis de ventana deslizante (*sliding window*) usando el software *PLINK*. Brevemente, este análisis busca una región para la cual los haplotipos tienen la asociación más fuerte con un rasgo binario, deslizando una ventana de una amplitud predeterminada sobre cada locus marcador y considerando todas las longitudes de haplotipos dentro de cada ventana. Para ilustrarlo consideremos una colección de datos de 10 loci. Cuando colocamos una ventana de 3 sobre el locus 5, las longitudes posibles de haplotipos que contienen el locus 5 son tres (loci 3-4-5, 4-5-6 y 5-6-7), dos (loci 4-5 y 5-6) y uno (locus 5). Para cada uno de estos subgrupos de loci se calcula un *score* estadístico que se basa en las diferencias entre el vector medio de los contajes de haplotipos para casos y controles. El máximo de estos *scores* estadísticos sobre todas las posibles longitudes de los haplotipos dentro de una ventana es la prueba estadística específica de locus. La prueba estadística global es el máximo sobre todos los *scores* estadísticos calculados. El estatus caso/control se permutó al azar mediante múltiples correcciones basadas en el número de SNPs correlacionados analizados para corregir los valores de P.

9.4. Estudio de interacciones

En el estudio de asociación del gen *LILRA3* se realizó un estudio para comprobar la posible interacción epistática entre este marcador y el principal alelo de susceptibilidad a EM, el *HLA-DRB1*15:01*. Este mismo análisis se realizó considerando la variable sexo, para observar si en nuestro estudio replicábamos las asociaciones previamente descritas en la

literatura^{140, 173}. El estudio de interacción entre marcadores de genes no ligados y el posterior análisis por sexo se realizaron mediante recuento directo a partir de tablas estratificadas.

9.5. Análisis de muestras procedentes de distintas poblaciones

Un meta-análisis se define como un análisis estadístico de una colección de resultados analíticos que se realiza con el propósito de integrar los distintos resultados observados¹⁷⁴ y proporciona una estrategia eficiente y práctica para detectar variantes con efecto pequeño, así como ofrece una estimación más precisa del efecto promedio de esa variante. El meta-análisis permite dar mayor “peso” a los estudios que presentan mayor número de casos/controles, mostrando el efecto individual de cada estudio así como el efecto combinado.

Los datos de muestras de distinto origen para los estudios presentados por Varadé *et al.*¹⁶⁰ y Ortiz *et al.*¹⁶³ fueron analizados con el programa *Review Manager 5.0 (RevMan 5.0. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaborations, 2008)*, que permite la realización del meta-análisis. La intensidad de la asociación se estudió con los valores de ORs en las colecciones individuales. El programa compara las distintas ORs e ICs para cada colección de datos mediante pruebas de homogeneidad implementadas en el programa (estadísticos χ^2 e I^2) para analizar si un factor tiene un efecto diferencial en las distintas poblaciones. Si el valor de P es superior al 10% (P de homogeneidad > 0,1), se asume que no hay heterogeneidad entre las distintas poblaciones. El estadístico I^2 indica la proporción de variación entre estudios respecto a la variación total, es decir, la proporción de variación total que es atribuible a la heterogeneidad. Aquellas poblaciones que mostraron una I^2 superior al 40% para un suceso determinado, se consideraron heterogéneas y en estos casos, se realizó un análisis de sensibilidad¹⁷⁴ eliminando secuencialmente cada cohorte para descartar la fuente de heterogeneidad. Finalmente, la prueba de asociación de *Mantel-Haenszel* ofrece los resultados combinados (P_{M-H} ; OR_{M-H}) entre poblaciones homogéneas, asumiendo como significativos aquellos valores de $P < 0,05$.

Para el estudio de los SNPs del gen *APOE*, el meta-análisis se llevó a cabo en “R” usando los paquetes estadísticos *HardyWeinberg v.1.4* y *rmeta v.2.16*. En este caso, los valores de P fueron de dos colas y la significación estadística a nivel del genoma completo se

definió como $P < 5 \times 10^{-8}$, mientras que la tendencia a la asociación se estableció en $P < 1 \times 10^{-4}$.

Para los estudios de cribado de SNPs y los estudios del gen *CD6*, la OR media procedente de la combinación de distintas colecciones de casos y controles, se calculó mediante una prueba de *Cochran-Mantel-Haenszel* implementado en *PLINK*. Este test considera cada conjunto de datos específico de una población como un estrato distinto para estudiar la asociación de los marcadores. La heterogeneidad de las ORs se estudió a través del test de *Breslow-Day*.

Finalmente, para el estudio de señales que quedaron al borde de la significación estadística del GWAS del 2011¹²⁰ se utilizó el software *METAL*¹⁷⁵. En primer lugar, *METAL* selecciona para cada marcador un alelo de referencia y se calcula un estadístico z que caracteriza la evidencia de asociación. El estadístico z resume la magnitud y la dirección del efecto del alelo de referencia y todos los estudios se alinean al mismo alelo de referencia. Luego, se calculan un estadístico z y un valor de P total mediante la suma ponderada de los estadísticos individuales. Los pesos son proporcionales a la raíz cuadrada del número de individuos estudiados en cada colección de datos y se seleccionan de tal forma que la suma de los pesos al cuadrado es 1.

En función del estudio, el meta-análisis que se realizó se basó en un modelo de efectos fijos o en un modelo de efectos aleatorios. Se optó por un modelo de efectos fijos cuando se asumió que las diferencias entre los tamaños de los efectos de las cohortes individuales se debían tan solo al hecho de que estas utilizan sujetos diferentes. Sin embargo, se optó por un modelo de efectos aleatorios cuando se asumía que podía haber efectos subyacentes que diferían entre las diferentes colecciones, teniendo en cuenta la heterogeneidad estimada entre las mismas. Finalmente, se realizaron estudios de sensibilidad para analizar la influencia de cada una de las colecciones en la estimación global del efecto para cada uno de los estudios presentados en este trabajo y evaluar así la robustez o estabilidad de la medida final obtenida.

El análisis de regresión logística es un modelo de regresión que permite estudiar si una variable binomial depende, o no, de otra u otras variables no necesariamente binomiales. Empleamos este tipo de análisis para estudiar el efecto que ejercen factores como la edad de debut o el sexo sobre los SNPs estudiados. En concreto se empleo en el estudio de los cinco genes sugerentes de asociación tras el GWAS llevado a cabo por Sawcer *et al.*¹²⁰ y en el

estudio del gen *APOE*¹⁵⁹. De esta forma se obtuvieron valores de ORs ajustadas a las variables mencionadas. El análisis de regresión logística se realizó en *PLINK*.

9.6. Análisis estadístico de los datos del FACS

Para calcular las diferencias en la expresión de moléculas de superficie se usaron pruebas de t de *Student* y ANOVA de dos colas según se comparasen dos o más factores, respectivamente. Las diferencias en la proliferación y la producción de citoquinas entre los diferentes haplotipos se estudió por la prueba no paramétrica U de Mann-Whitney usando el software *Graphpad Prism v.5* (*GraphPad Software, La Jolla, CA*).

Resultados y discusión

1. Estudio de cribado de genes de candidatos y validación de las señales significativas

La identificación inequívoca por medio de los GWAS de loci de susceptibilidad, tales como *IL7R*, *IL2RA* e *IL12A*, ha certificado a los genes de citoquinas y de receptores de las mismas como una clase genuina de loci de riesgo a EM^{111, 176-178}. Por lo tanto, el presente estudio¹⁶² se diseñó con el propósito primario de investigar a través de una aproximación *haptag*, aquellos genes de citoquinas, sus receptores y factores de transducción de la señal asociados, que no habían sido estudiados de forma sistemática en estudios de gen candidato previos a los GWAS en EM. Por lo tanto, se excluyeron del presente estudio genes de loci tales como *IFNG*, *IL4*, *IL1A/IL1B* que habían sido estudiados anteriormente¹⁷⁹. Además, en el actual cribado se incluyó una selección de genes codificantes para receptores ionotrópicos y transportadores de glutamato, que regulan los niveles de glutamato cerebrales ya que altos niveles de este neurotransmisor están implicados en mecanismos patológicos que conducen a la muerte neuronal. También se ha descrito que los niveles de glutamato en los cerebros de pacientes con EM están al menos parcialmente controlados por polimorfismos genéticos comunes, que SNPs de genes que pertenecen al sistema glutamatérgico están asociados con la capacidad de respuesta a la terapia con IFN- β en EM¹⁸⁰⁻¹⁸² y que el daño glial mediado por glutamato contribuye a la patología de la materia blanca, como se ha observado en EM¹⁸³.

El estudio se llevó a cabo en dos fases. En primer lugar, se realizó un cribado con 384 SNPs en la colección de Bilbao, de los cuales 368 SNPs se genotiparon satisfactoriamente para 462 casos y 470 controles (eficacia media de genotipado: 97,3%). Los alelos de riesgo y sus frecuencias, los valores de P para la asociación con la enfermedad y las ORs de los 20 SNPs más fuertemente asociados ($P \leq 0,021$) derivados del cribado inicial se muestran en la Tabla 9.

Tabla 9. Resultados de los veinte SNPs más fuertemente asociados ($P \leq 0,02$) en el cribado inicial en la cohorte de Bilbao.

Gen	SNP ^a	% Alelo de riesgo		P	OR (IC 95%)
		Casos	Controles		
<i>SOCS1</i> (5'-UTR)	rs243324 (C/ <u>T</u>)	0,55	0,47	0,0005	1,38 (1,18–1,62)
<i>IL28RA</i> (4 ^{to} intrón)	rs1416834 (A/ <u>G</u>)	0,41	0,33	0,0008	1,38 (1,17–1,63)
<i>OSMR</i> (1 ^{er} intrón)	rs3805558 (<u>G</u> /A)	0,75	0,69	0,0017	1,38 (1,16–1,65)
<i>IL7R</i> (6 ^{to} intrón)	rs6897932 (<u>C</u> /T)	0,76	0,7	0,0022	1,38 (1,15–1,66)
<i>IL22RA2</i> (5 ^{to} intrón)	rs202573 (G/ <u>A</u>)	0,33	0,27	0,0026	1,36 (1,14–1,62)
<i>IL28RA</i> (2 ^{do} intrón)	rs7520329 (C/ <u>A</u>)	0,3	0,24	0,0026	1,37 (1,14–1,64)
<i>GRIK3</i> (6 ^{to} intrón)	rs12067006 (T/ <u>C</u>)	0,29	0,22	0,0034	1,37 (1,14–1,66)
<i>IL7R</i> (5'-UTR)	rs6890853 (<u>G</u> /A)	0,75	0,69	0,0047	1,34 (1,12–1,61)
<i>IL31</i> (3'-UTR)	rs7310689 (<u>G</u> /A)	0,95	0,91	0,0067	1,66 (1,21–2,28)
<i>OSMR</i> (7 ^{mo} intrón)	rs12657342 (<u>C</u> /G)	0,79	0,74	0,013	1,31 (1,09–1,59)
<i>IL32</i> (3'-UTR)	rs2015620 (T/ <u>A</u>)	0,34	0,29	0,015	1,27 (1,07–1,51)
<i>OSMR</i> (3 ^{er} intrón)	rs2278324 (C/ <u>A</u>)	0,15	0,11	0,015	1,40 (1,10–1,79)
<i>IL1RL2</i> (8 ^{vo} intrón)	rs17637748 (<u>A</u> /G)	0,57	0,52	0,016	1,25 (1,07–1,47)
<i>LIF</i> (5'-UTR)	rs3761427 (<u>C</u> /G)	0,65	0,6	0,017	1,26 (1,07–1,49)
<i>IL1RL2</i> (11 ^{er} intrón)	rs2287040 (<u>G</u> /A)	0,68	0,63	0,018	1,26 (1,06–1,49)
<i>OSMR</i> (9 ^{no} intrón)	rs10040172 (A/ <u>G</u>)	0,17	0,13	0,018	1,36 (1,08–1,70)
<i>OSM</i> (3'-UTR)	rs11089441 (<u>G</u> /T)	0,64	0,58	0,018	1,25 (1,06–1,48)
<i>IL15</i> (1 ^{er} intrón)	rs1589241 (<u>C</u> /T)	0,68	0,63	0,019	1,26 (1,06–1,49)
<i>IL17B</i> (1 ^{er} intrón)	rs353268 (G/ <u>A</u>)	0,09	0,06	0,019	1,52 (1,12–2,06)
<i>IL1F6</i> (3'-UTR)	rs1867828 (G/ <u>A</u>)	0,25	0,2	0,021	1,30 (1,07–1,57)

^a Los alelos de riesgo están subrayados. IC: intervalo de confianza; OR: *odds ratio*; SNP: polimorfismo de un solo nucleótido; UTR: *untranslated region*.

Los cinco SNPs más fuertemente asociados se encontraron dentro o cerca de los genes codificantes para el gen supresor de la señalización de citoquinas 1 (*SOCS1*), la cadena α del receptor de la IL-28 (*IL28RA*), el receptor de la oncostatina M (*OSMR*), el receptor de la IL-7 (*IL7R*) y el receptor $\alpha 2$ de la IL-22 (*IL22RA2*), con ORs entre 1,36-1,38 y valores de P no corregidos comprendidos entre 0,003-0,0005 (ninguna de estas asociaciones soportó la *corrección de Bonferroni*) (Tabla 9). Dentro de esta primera fase, se estudiaron 7 SNPs del gen *IL7R* y entre ellos el rs6897932 emergió como aquel con una asociación más fuerte, en concordancia con estudios anteriores^{111, 177, 178}. La detección de este SNP de riesgo ya

establecido puede dar fe de la idoneidad de esta colección de casos y controles para descubrir nuevos loci de riesgo con OR similares o más altas. Posteriormente para la corrección por hipótesis múltiples llevamos a cabo un test de permutación como alternativa a la *corrección de Bonferroni* y se observó que *SOCS1* (rs243324: $P = 0,026$) e *IL28RA* (rs1416834: $P = 0,05$) eran los loci más fuertemente asociados ($P_{\text{corregida por 10.000 permutaciones}} < 0,05$). Sin embargo, los cuatro SNPs más fuertemente asociados que no se habían descrito anteriormente (el rs6897932 de *IL7R* se excluyó) fueron analizados posteriormente en una cohorte de validación que incluía cinco colecciones de muestras independientes de origen europeo y norteamericano, con un total de 3.919 pacientes con EM y 4.003 controles (Tabla 10).

Tabla 10. Características clínicas y demográficas de los controles y los pacientes usados en el estudio.

Población	Casos			Controles	
	n	Hombre/Mujer/ n.d. (%)	RR y SP/PP/ Otras/n.d. (%)	n	Hombre/Mujer/ n.d. (%)
Bilbao	462	29/71/0	81/10/2/7	470	28/71/1
Barcelona	708	37/63/0	78/21/1	825	46/53/1
Andalucía	1.131	30/60/10	80/1/1/18	1.237	30/68/2
Madrid	680	34/64/2	83/10/1/6	832	45/53/2
UCSF Afroamericanos	895	21/77/2	82/7/7/4	658	29/70/1
UCSF Caucásicos	505	32/68/0	80/3/17/0	451	33/67/0

n.d.: no disponible; PP: primaria progresiva; RR: remitente recurrente; SP: secundaria progresiva; UCSF: Universidad de California, San Francisco.

La Tabla 11 resume los resultados de los análisis de validación y los análisis combinados. El rs6897932 en *IL7R* se excluyó del estudio de validación dada su asociación bien establecida con EM^{111, 177, 178}. En este estudio¹⁶², el rs1416834 en el gen *IL28RA* mostró un patrón de asociación del alelo G en Barcelona [$P = 0,05$; OR (IC 95%) = 0,86 (0,75-1,00)] y Madrid [$P = 0,004$; OR (IC 95%) = 0,79 (0,68-0,93)] opuesto al observado en la colección de Bilbao [$P = 0,0008$; OR (IC 95%) = 1,38 (1,17-1,63)]. Debido a este efecto de heterogeneidad, el test de *Breslow-Day* para el rs141684 combinando la información de las colecciones primarias y las de validación fue altamente significativo ($P = 0,0002$). El

rs3805558 del gen *OMSR* no mostró una asociación significativa en las colecciones de validación ni de forma combinada. El alelo A del rs202573 en el gen *IL22RA2* se asoció significativamente en la colección de Bilbao [P = 0,0026; OR (IC 95%) = 1,36 (1,14-1,62)] y Andalucía [P=0,04; OR (IC 95%) = 1,14 (1,00-1,29)], pero no fue significativo en ninguna de las colecciones de validación restantes ni de forma combinada.

Tabla 11. Estudio de asociación de los 4 SNPs más asociados en el cribado inicial en las colecciones de validación y análisis combinado.

Gen	SNP	Cohorte de validación			Combinado ^a	
		Alelo de riesgo	P _{CMH}	OR (IC 95%)	P _{CMH}	OR (IC 95%)
<i>SOCS1</i>	rs243324	T	0,0027	1,10 (1,03-1,18)	0,00006	1,13 (1,07–1,20)
<i>IL28RA</i>	rs1416834	A ^b	0,06	1,06 (1,00-1,14)	- ^c	- ^c
<i>OSMR</i>	rs3805558	G	0,52	1,02 (0,95-1,10)	0,09	1,06 (0,99–1,13)
<i>IL22RA2</i> ^d	rs202573	A	0,42	1,03 (0,96-1,11)	0,06	1,07 (1,00–1,15)

IC: intervalo de confianza; OR: *odds ratio*; SNP: polimorfismo de un solo nucleótido. ^a Colecciones de Bilbao y de validación combinadas; ^b el alelo de riesgo en la cohorte de validación es opuesto al identificado en la colección de Bilbao (Tabla 9); ^c test de *Breslow-Day* para la heterogeneidad del efecto entre las poblaciones fue significativo (P = 0,0002); ^d se excluyó la colección afroamericana por una desviación del HWE en los controles (P = 0,0006).

Finalmente, el rs243324 localizado en la región reguladora 5' del *SOCS1* (*suppressor of cytokine signaling-1*) mostró ser el único SNP significativamente asociado en la cohorte de validación (P_{CMH} de validación = 0,0027). Así mismo, del estudio combinado de los datos del cribado primario y de los datos de validación en más de 4.300 pacientes con EM y 4.400 controles, el rs243324 destacó por estar significativamente asociado [P = 6,01x10⁻⁵; OR (IC 95%) = 1,13 (1,07-1,20)] y emergió como uno de los nuevos marcadores de susceptibilidad a EM. De forma similar a lo observado en Bilbao, el alelo T del rs243324 era el alelo de riesgo asociado significativamente en la colección de Madrid [P = 0,0004; OR (IC 95%) = 1,30 (1,12-1,51)] y Andalucía [P = 0,045; OR (IC 95%) = 1,13 (IC 95%) = 1,00-1,27)], pero no se encontraron valores significativos en las tres colecciones restantes (datos no mostrados; P_{Breslow-Day} = 0,14 para la cohorte de validación completa).

A partir de este punto, se tuvieron en cuenta cuatro consideraciones en relación a *SOCS1*. En primer lugar, dado que la arquitectura del LD varía entre poblaciones de ascendencia europea o africana, también analizamos los patrones de asociación de los SNPs estudiados en la fase de validación excluyendo la colección de datos afroamericana. Esto incrementó la fuerza de la asociación para el rs243324 en *SOCS1* en la cohorte total de validación [$P_{\text{CHM}} = 0,0005$; OR (IC 95%) =1,16 (1,07-1,26)] y de forma combinada incluyendo la colección de Bilbao [$P_{\text{CHM}} = 0,00004$; OR (IC 95%) = 1,19 (1,11-1,29)]. Sin embargo, al realizar esta exclusión no se observaron diferencias en los patrones de asociación de los restantes SNPs estudiados en esta fase.

En segundo lugar, tuvimos en consideración que el gen *SOCS1* se localiza en el cromosoma 16 a una distancia de alrededor de 70 kb de *CLEC16A*, que es un gen de riesgo confirmado para EM^{111, 116, 184}. Por lo tanto, decidimos investigar el LD entre una selección de SNPs en *CLEC16A* y *SOCS1*. Los SNPs estudiados fueron: rs12708716 o su equivalente rs2903692, rs6498169 y rs243324. Este análisis se realizó solo en las colecciones de Bilbao, Madrid y Andalucía (Tabla 12).

Tabla 12. Estudio de asociación de SNPs en *CLEC16A* y desequilibrio de ligamiento con *SOCS1*.

Colección de datos	<i>CLEC16A</i>	Posición	Alelo de riesgo; % EM / % Controles	P	OR (IC 95%)	D'/r ² con rs243324
Bilbao	rs12708716 ^a	11087374	A; 0,64/0,61	0,24	1,11 (0,93–1,31)	0,29/0,052
	rs6498169	11156830	G; 0,37/0,35	0,25	1,10 (0,93–1,31)	0,23/0,030
Madrid	rs2903692 ^a	11146284	G; 0,66/0,59	0,0002	1,35 (1,15–1,58)	0,26/0,047
	rs6498169	11156830	G; 0,39/0,32	0,0007	1,35 (1,13–1,61)	0,25/0,029
Andalucía	rs2903692 ^a	11146284	G; 0,65/0,60	0,0003	1,26 (1,11–1,42)	0,31/0,063
	rs6498169	11156830	G; 0,36/0,34	0,095	1,11 (0,98–1,26)	0,27/0,034

IC: intervalo de confianza; OR: *odds ratio*; SNP: polimorfismo de un solo nucleótido. ^a El rs12708716 y el rs2903692 son equivalentes (D'=1, r²=0,96; <http://www.broadinstitute.org/mpg/snap/>).

Ninguno de los SNPs del *CLEC16A* estaba significativamente asociado con EM en la colección de Bilbao, pero el rs2903692 mostró una asociación en las colecciones de Madrid y Andalucía. Los análisis de los patrones de LD de cada uno de los SNPs del *CLEC16A* con el rs243324 del *SOCS1*, confirmaron que *SOCS1* es un locus de riesgo independiente para EM ($D' = 0,25-0,31$; $r^2 = 0,029-0,063$) en cada una de las tres colecciones. Esto se debe probablemente a la presencia de un área que muestra una tasa de recombinación de hasta 8 cM/Mb y que separa los bloques de LD de *CLEC16A* y *SOCS1*.

En tercer lugar, consideramos los distintos estudios que se habían realizado en relación a *SOCS1* o la proteína que codifica y EM. Se ha visto que SOCS1 funciona impidiendo la vía de señalización de IFNs tipo I y II por unirse predominantemente a IFNAR1 fosforilada o proteínas JAK asociadas a IFNGR a través de su dominio SH2, anulando la activación STAT y la subsiguiente expresión de entre otros, los genes del HLA de clase I y II y el gen *CD40*¹⁸⁵. Un papel para SOCS1 en condiciones desmielinizantes se ha inferido principalmente de estudios con el modelo experimental EAE. La expresión del mRNA SOCS1 en EAE se restringe principalmente a células mononucleares infiltrantes¹⁸⁶, pero los niveles de mRNA de SOCS1 de médula espinal son más altos en el pico de la EAE crónica (C57BL/6J; inducida por glicoproteínas oligodendrocitarias de la mielina) comparados con la EAE-RR (SJL/J; inducidos por mielina bovina)¹⁸⁷. Los ratones transgénicos que sobreexpresan SOCS1 en oligodendrocitos desarrollan EAE en menor tiempo acompañada por una inflamación temprana mayor; un efecto relacionado con la falta de respuesta de los oligodendrocitos a los efectos protectores tempranos del IFN- γ ¹⁸⁸. La administración de SOCS1 mimético, un péptido inhibidor de la tirosina kinasa, suprimió el desarrollo de la EAE en la cepa de ratones New Zealand White, protegió a los ratones SJL/J con EAE-RR contra los brotes¹⁸⁹ y redujo la severidad de la enfermedad en ratones C57BL/6 con EAE crónica¹⁹⁰. Los ratones deficientes en SOCS1 específico de LT son esencialmente resistentes a EAE debido a la diferenciación preferencial de LTs CD4⁺ hacia Th1 más que hacia Th17, el tipo Th esencial para EAE¹⁹¹. Berard *et al.*¹⁹⁰ analizaron la expresión de SOCS1 en las formas crónicas y RR de la EAE, inducida en la misma cepa de ratón (C57BL/6) usando el mismo antígeno de la mielina (glicoproteína oligodendrocitaria de la mielina). En ambos modelos, SOCS1 se expresó predominantemente en macrófagos. La proporción de macrófagos Mac-1⁺ que expresan SOCS1 en las lesiones en el pico de la EMRR fue significativamente más alto que en las observadas en EAE crónica y coincidió con una

menor expresión de las moléculas efectoras de macrófagos inducibles por la óxido nítrico sintasa ¹⁹⁰. En conjunto, estos datos indican que la expresión aumentada de *SOCS1* en macrófagos en el pico de la EMRR puede conducir a la desconexión de rutas efectoras proinflamatorias promoviendo así la remisión en el modelo de EAE-RR. Teniendo en cuenta estos datos, decidimos estudiar si las frecuencias alélicas para el rs243324 del *SOCS1* diferían entre los pacientes de las formas EMPP y los de las formas EMRR/EMSP. A través del presente estudio ¹⁶², que no fue *a priori* diseñado para estudiar el impacto de la variación genética en el curso clínico de la EM, el número total de pacientes disponibles de las formas EMPP (n = 347) y de las formas EMRR/EMSP (n = 3.545) en las seis colecciones de individuos (Tabla 10) era suficientemente alto como para alcanzar una potencia comprendida entre 73 y 81% para detectar los efectos genéticos conferidos por el rs243324 con una OR de 1,25, bajo modelos aditivos o multiplicativos, respectivamente.

La Tabla 13 muestra que la frecuencia del alelo T es de media un $5,2 \pm 1,9\%$ más alta en las formas EMRR/EMSP que en las EMPP. Este efecto fue homogéneo en las 6 colecciones de datos incluidas en el presente estudio ($P_{\text{Breslow-Day}} = 0,96$). Debido al relativamente bajo número de formas EMPP en cada estrato, no se observó significación en ninguna colección de forma individual, sin embargo, el efecto combinado de todas las colecciones juntas sí fue estadísticamente significativo [$P_{\text{CMH}} = 0,0096$; OR (IC 95%) = 1,24 (1,05-1,46)].

En cuarto y último lugar, en relación al rs243324 del *SOCS1* se realizaron análisis estratificados de las colecciones en función del género de los individuos. Los análisis individuales y combinados revelaron la ausencia de cualquier efecto significativo debido a esta variable (datos no mostrados).

Así, aunque se requiere una confirmación independiente para validar firmemente estos descubrimientos, y estudios funcionales adicionales para evaluar el impacto del rs243324 en la expresión de *SOCS1*, el locus *SOCS1* parece emerger de este estudio como un marcador potencial del curso clínico de la EM. Además, en otros estudios se ha encontrado que algunos polimorfismos funcionales entre los que se incluye el rs33977706, alteran la actividad transcripcional del gen *SOCS1* ^{192, 193}.

Tabla 13. Frecuencias alélicas del rs243324T del gen *SOCS1* en pacientes con forma RR/SP y PP.

Colección de datos	RR/SP	PP	P	OR (IC 95%)
UCSF Afroamericanos	0,48	0,45	0,61	1,10 (0,76–1,59)
UCSF Caucásicos	0,53	0,50	0,76	1,12 (0,55–2,26)
Bilbao	0,56	0,48	0,14	1,41 (0,89–2,21)
Barcelona	0,56	0,50	0,09	1,25 (0,96–1,61)
Andalucía	0,56	0,50	0,58	1,27 (0,55–2,94)
Madrid	0,58	0,53	0,29	1,23 (0,84–1,80)
Combinado	0,54	0,49	0,0096 ^a	1,24 (1,05–1,46) ^a

IC: intervalo de confianza; OR: *odds ratio*; PP: primaria progresiva; RR: remitente recurrente; SP: secundaria progresiva; UCSF: Universidad de California, San Francisco. ^a Test de Cochran-Mantel-Haenszel, test de heterogeneidad de *Breslow-Day*, $P = 0,96$.

Un estudio reciente analizó genes relacionados funcionalmente con el locus *IL7RA*, validado para riesgo a EM, e identificó una serie de SNPs *upstream* del *SOCS1* asociados con la susceptibilidad a EM ¹⁹⁴. El SNP rs441349 fue el más significativamente asociado en este estudio y se localiza en el área intergénica entre protamina-2 (*PRM2*) y *PRM1* a una distancia de 16,8 kb centromérica del rs243324 y a 21,8 kb del sitio de inicio de la transcripción de *SOCS1* ¹⁹⁴. El rs441349 se genotipó en las cohortes de Madrid y Andalucía (datos no mostrados), y se encontró que de forma semejante a Zuvich *et al.* ¹⁹⁴, el alelo C constituía el alelo de riesgo asociado con EM ($P_{CMH} = 0,01$). La significación de la asociación del SNP rs441349 fue menor que la del rs243324 ($P_{CMH} = 0,0002$) cuando ambos SNPs se consideraron juntos en las colecciones españolas. El LD entre ambos SNPs es bajo ($D' = 0,765$; $r^2 = 0,11$). El análisis de haplotipos mostró que el haplotipo (CT) formado por los SNPs rs243324–rs441349 describe la asociación con EM marginalmente mejor que el rs243324 individualmente, pero mucho mejor que el rs441349 de manera individual [P corregida por 10.000 permutado = 0,0001 (haplotipo CT), 0,0004 (rs243324) y 0,026 (rs441349)]. Actualmente se está llevando a cabo un extenso mapeo fino del área del gen *SOCS1* incluyendo un grupo de cuatro pequeños genes compuestos de *PRM1-3* y proteína de transición-2. Las protaminas y las proteínas de transición-2 facilitan la condensación de material genético dentro de la espermátida en desarrollo ¹⁹⁵. Aunque parecen haberse publicado asociaciones no funcionales con autoinmunidad, inflamación y EM, la expresión

coordinada de este locus multigénico ¹⁹⁵ puede justificar una investigación conjunta en estudios funcionales centrados en *SOCS1*.

Otro de los genes estudiados en la fase de validación, pero sin alcanzar un resultado significativo es el gen *IL22RA2*. El gen *IL22RA2* codifica un receptor secretado soluble de IL-22. El rs276474, localizado a una distancia de 8,9 kb en el extremo 3' del gen *IL22RA2*, se encontró recientemente asociado con EM en una cohorte combinada de Suecia y Noruega ¹⁹⁶. En nuestro estudio ¹⁶², el rs202573, localizado en el quinto intrón del *IL22RA2*, se asoció en las cohortes individuales de Bilbao y Andalucía, pero su asociación no fue significativa en la cohorte total. Ambos SNPs se localizan a 17,6 kb de distancia el uno del otro y apenas presentan desequilibrio de ligamiento entre ellos ($r^2 = 0,013$, $D' = 0,20$; base de datos del “Proyecto 1.000 Genomas”).

El rs1416834 que se localiza en el cuarto intrón del *IL28RA*, mostró un patrón alélico de asociación en la cohorte de Bilbao opuesto al que se observa en las cohortes de Madrid y Barcelona. Mientras que en la cohorte de Bilbao el alelo G del SNP rs1416834 es el que confiere riesgo a padecer EM [$P = 0,008$; OR (IC 95%) = 1,38 (1,17-1,63)] en la cohorte de validación es el alelo A el que parece confiere el riesgo [$P = 0,06$; OR (IC 95%) = 1,06 (1,00-1,14)]. En un GWAS reciente se identificó el rs4649203 en la región promotora del *IL28RA* como un marcador de susceptibilidad para psoriasis ¹⁹⁷. El rs1416834 y el rs4649203 se separan por un intervalo de 33 kb que contiene un punto de recombinación de 15 cM/Mb alrededor del segundo exón y por lo tanto, no están en LD ($r^2 = 0,033$; $D' = 0,28$; base de datos del “Proyecto 1.000 Genomas”). Aunque es posible que la asociación del rs1416834 del *IL28RA* con EM represente un descubrimiento espurio, no se puede excluir que este patrón de asociación alélica diferencial entre cohortes con el riesgo al padecer EM, denominado ‘flip-flop’, sea debido a una variación muestral o poblacional en correlaciones interlocus con la variante causal real ^{198, 199}. Dado que la EM y la psoriasis comparten el locus de susceptibilidad TYK2 ^{134, 194, 197}, es necesario realizar un mapeo fino de la región *IL28RA* para conocer si este locus representa un segundo determinante genético compartido.

2. Estudio de asociación a gran escala del gen *APOE* en esclerosis múltiple

La investigación del gen *APOE*, el locus de riesgo más importante para la enfermedad de Alzheimer, ha sido motivada en EM por publicaciones de ligamiento genético de la región que contiene este gen en el cromosoma 19q13 a EM⁹⁵, como también, por su papel funcional en el transporte de lípidos, la inmunorregulación, la neuroplasticidad y los mecanismos reparadores⁹⁶. Sin embargo, los estudios de asociación de *APOE* con EM habían conducido a resultados principalmente negativos, con algunos estudios describiendo efectos significativos, mientras muchos otros eran incapaces de confirmar estas asociaciones. En este estudio nos propusimos evaluar de forma concluyente la asociación entre los dos SNPs más comúnmente estudiados del gen *APOE* y la susceptibilidad a EM en poblaciones de ascendencia europea. Para este objetivo, se llevó a cabo un estudio de asociación a gran escala de los SNPs rs429358 y rs7412 con el riesgo a padecer EM.

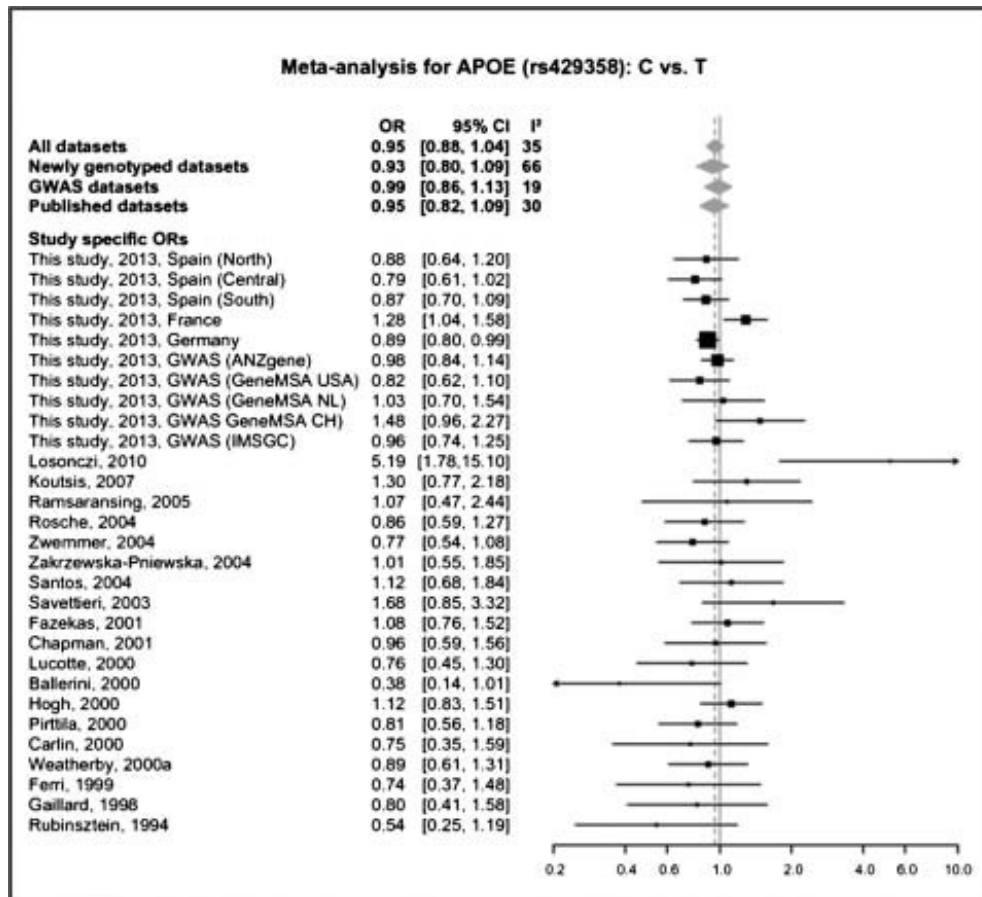
Para este fin, se emplearon muestras de tres fuentes distintas. En primer lugar, se genotiparon muestras de 12.740 sujetos no estudiados hasta entonces para estos dos SNPs (6.741 casos y 5.999 controles). Posteriormente, se imputaron datos de genotipos de 8.739 sujetos de las colecciones independientes de 5 GWAS usando los paneles de referencia de alta resolución más recientes (3.530 casos y 5.209 controles); y en último lugar, se recopilaron datos de genotipos de estudios previos de 8.265 (3.642 casos y 4.623 controles) y 6.981 (2.931 casos y 4.050 controles) sujetos para el rs425893 y el rs7412, respectivamente.

En total, nuestro estudio que incluye 29.744 (13.913 casos y 15.831 controles) y 28.460 (13.202 casos y 15.258 controles) sujetos para el rs429358 y el rs7412, respectivamente, superó de forma destacada en número, a meta-análisis anteriores de estos 2 SNPs en los que se combinaban los datos previamente publicados y en los que se incluían 5.831²⁰⁰ y 7.706 sujetos de origen caucásico²⁰¹. Además en los análisis combinados, nuestro estudio tenía una potencia estadística de aproximadamente un 99% y un 72% para el rs429358 y el rs7412, respectivamente, para detectar una OR de 1,20 con un umbral de significación a nivel del genoma completo ($\alpha = 5 \times 10^{-8}$).

El análisis de nuestros datos no mostró una evidencia de asociación de ninguno de estos dos SNPs del gen *APOE* con el riesgo a padecer EM [rs429358: P=0,259; OR (IC 95%)=0,95 (0,88-1,04); y rs7412: P=0,0569; OR (IC 95%)=1,07 (1,00-1,15); Figura 11] con

una moderada evidencia de heterogeneidad para el rs429358 ($I^2=35$) y sin heterogeneidad para el rs7412 ($I^2=0$).

(A)



(B)

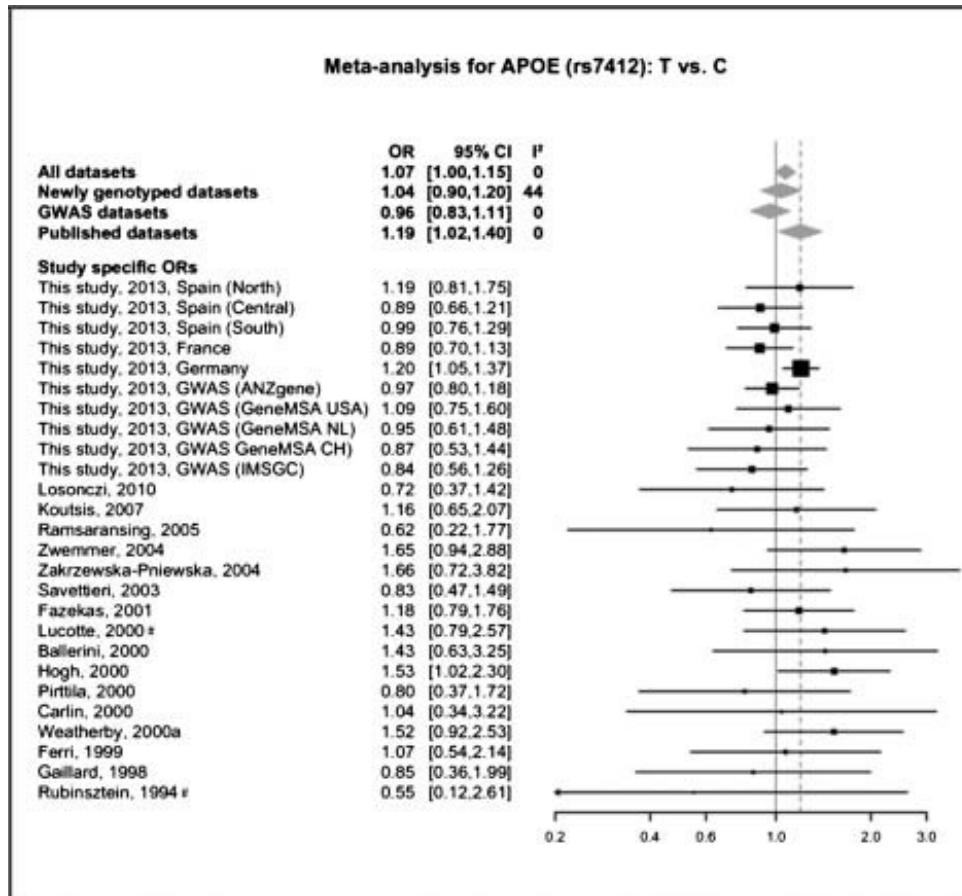


Figura 11. Forest plot de los meta-análisis realizados estudiando la asociación entre los SNPs rs429358 (A) y el rs7412 (B) y el riesgo a EM en poblaciones de ascendencia europea. Las ORs específicas de estudio (cuadrados negros) y los ICs al 95% (líneas) se calcularon para cada colección imputada. La OR total y los ICs al 95% (diamantes grises) se calcularon combinando todas las colecciones y de forma individual para cada una de las procedencias de las colecciones. # La exclusión de los estudios que no se encontraban en HWE no cambió sustancialmente los resultados del meta-análisis (datos no mostrados). ANZgene: Consorcio sobre Genética de la Esclerosis Múltiple de Australia y Nueva Zelanda, APOE: gen de la apolipoproteína E; CH: Suiza, NL: Países Bajos, GWAS, estudio de asociación mediante barrido genómico, IMSGC: Consorcio Internacional sobre Genética de la Esclerosis Múltiple, MS: esclerosis múltiple.

El análisis estratificado para el SNP rs429358 agrupando las colecciones de datos en función de la fuente de la que procedían mostró resultados similares a los observados antes de la estratificación (Figura 11 A). En el meta-análisis de las colecciones que habían sido previamente publicadas para el SNP rs7412 se observó una OR relativamente pronunciada de 1,19 ($P = 0,0263$) sin evidencia de heterogeneidad ($I^2 = 0$, Figura 11 B) y además, se observó que el test de regresión también mostraba evidencias significativas de efectos de pequeño tamaño ($P = 0,0047$). Esto puede indicar la presencia de un sesgo de publicación o un informe selectivo. En este contexto, hay que tener en cuenta que dos publicaciones incluían datos relativos al rs429358 pero no al rs7412, aunque dichos datos se habían generado en estos estudios^{202, 203}. En esta línea, el meta-análisis del rs7412 estratificado tanto para las colecciones imputadas de los GWAS como para aquellas genotipadas para este estudio, mostró en ambos casos un efecto no significativo ($OR = 0,96$; $P = 0,579$ y $OR = 1,04$; $P = 0,616$, respectivamente, Figura 11 B).

Nuestro estudio, dado su tamaño muestral y su inherente potencia estadística, sugiere que son improbables incluso efectos modestos del rs429358 y del rs7412 en el riesgo a padecer EM. El hecho de que los estudios previos no hayan mostrado resultados concluyentes se puede atribuir probablemente a la falta de potencia, un problema que se supera con el presente análisis. Sin embargo, a pesar del gran tamaño muestral de nuestro estudio, no se puede excluir que el rs429358 y el rs7412 del *APOE* ejerzan efectos genéticos más pequeños que los asumidos aquí. Por ejemplo, en el reciente GWAS de EM¹²⁰, el tamaño del efecto que más pequeño que alcanzó una significación a nivel del genoma completo fue de 1,09, pero dadas las frecuencias alélicas del rs429358 y del rs7412 se requeriría un total de 88.000 y 140.000 sujetos, respectivamente, para alcanzar un 80% de potencia para detectar tales tamaños del efecto con un umbral de significación a nivel del genoma completo. Aunque nuestro estudio carecía de potencia para detectar tales efectos de menor tamaño, ni siquiera observamos una tendencia a la asociación en las colecciones actuales haciendo que la existencia de cualquier papel de los 2 SNPs investigados del *APOE* sea altamente improbable. Además, debido a la falta de acceso a los datos de genotipos a nivel individual y a los datos de las covariables, no fuimos capaces de ajustar los resultados de asociación de las colecciones anteriormente publicadas para la edad de diagnóstico y el sexo. Sin embargo, en aquellas colecciones donde fue posible la inclusión de estas covariables, los resultados no mostraron un efecto apreciable de la edad del diagnóstico y el

sexo con el riesgo a padecer la enfermedad. Además, solo estudiamos la asociación potencial de EM con las variantes funcionales rs429358 y rs7412 y no podemos excluir que otras variantes en esta región, que presenten nulo o débil LD con estos 2 SNPs, ejerzan un efecto en el riesgo a padecer EM. Sin embargo, esto último parece improbable debido a los resultados observados en los recientes GWAS^{111, 204} y en los meta-análisis de los mismos en EM^{116, 121}, que han capturado una fracción sustancial de la variación genética en este locus, excepto para los SNPs rs429358 y rs7412, que no se pudieron imputar efectivamente hasta muy recientemente.

Finalmente, nuestro estudio se restringió solamente a estudiar la influencia de los SNPs del gen *APOE* en la susceptibilidad a EM, pero no se puede excluir una posible influencia de estos SNPs en otros rasgos relevantes tales como la progresión o los cambios en las imágenes de resonancia magnética en esta enfermedad. Los estudios investigando la asociación de los SNPs del *APOE* y estas variables en EM son numerosos y muestran resultados contradictorios, posiblemente debido a los pequeños tamaños muestrales. Sin embargo, los estudios publicados hasta la fecha no han sugerido un efecto notable del *APOE* como predictor de la severidad a EM, la cognición o la atrofia²⁰⁵.

3. *Follow-up* del GWAS desarrollado por el IMSGC y estudio de otros genes candidatos

En este trabajo nos propusimos validar genes que mostraban una tendencia a la asociación como factores de riesgo a EM. Para ello incluimos 7 SNPs con ORs > 1,27 que no alcanzaron el umbral de significación ($P < 4,6 \times 10^{-4}$) en un estudio exhaustivo de replicación²⁰⁶ del primer GWAS llevado a cabo en EM¹¹¹, para su validación en una población independiente. Además estudiamos el gen *CBLB*, descrito como loci de riesgo en Cerdeña y posteriormente asociado solo en población italiana [$P = 1,60 \times 10^{-10}$ [OR (IC 95%) = 1,40 (1,27-1,57)]^{109, 110} para ver su efecto en nuestra población; y el gen *IL12B*, debido a los resultados observados en estudios de genes candidatos en dermatitis atópica¹⁰⁰, colitis ulcerosa¹⁰³, espondilitis anquilosante¹⁰⁴ y psoriasis¹⁰⁰⁻¹⁰². Finalmente, incluimos el gen *CNRI*, asociado anteriormente con la EMPP en estudios de genes candidatos⁹⁸, conscientes de la necesidad de replicación para discriminar verdaderas asociaciones de resultados espurios.

Se incluyeron 2.863 pacientes de EM y 2.930 controles sanos en este estudio¹⁶⁰. Aunque 4 de los 7 primeros SNPs mencionados en este estudio presentaban una MAF $\leq 5\%$, existía una potencia estadística mayor del 80% para detectar los efectos previamente observados (ORs > 1,27)²⁰⁶ (Tabla 14). El análisis de SNPs comunes en los GWAS (MAF >5%) no ha sido capaz de explicar la heredabilidad total estimada para enfermedades complejas como la EM y algunos autores especulan que parte de la heredabilidad perdida podría radicar en SNPs no comunes¹²⁸.

Tabla 14. Distribución de las frecuencias alélicas y genotípicas (n, %) de los 11 SNPs estudiados.

Gen	SNP	Genotipos	Casos	Controles	P*	OR (IC 95%)
		Alelos	n (%)	n (%)		
<i>PKN2</i>	rs305217	TT	4 (0,1)	6 (0,2)	0,424	0,92 (0,76-1,13)
		CT	209 (7,3)	226 (7,8)		
		CC	2.646 (92,5)	2.663 (92,0)		
		T	217 (3,8)	238 (4,1)	0,385	0,92 (0,76-1,11)
		C	5.501 (96,2)	5.552 (95,9)		
<i>GTF2B</i>	rs7538427	CC	3 (0,1)	9 (0,3)	0,082	0,84 (0,69-1,03)
		CT	192 (6,8)	224 (7,8)		
		TT	2.648 (93,1)	2.656 (91,9)		
		C	198 (3,5)	242 (4,2)	0,049	0,83 (0,68-1,00)
		T	5.488 (96,5)	5.536 (95,8)		
<i>EPHA4</i>	rs1517440	CC	5 (0,2)	7 (0,2)	0,479	1,07 (0,89-1,29)
		CT	252 (8,8)	238 (8,2)		
		TT	2.603 (91,0)	2.651 (91,5)		
		C	262 (4,6)	252 (4,4)	0,551	1,06 (0,88-1,26)
		T	5.458 (95,4)	5.540 (95,6)		
<i>TMEM39A</i>	rs1132200	AA	39 (1,4)	68 (2,4)	0,005	0,84 (0,75-0,95)
		AG	640 (22,4)	712 (24,6)		
		GG	2.180 (76,3)	2.112 (73,0)		
		A	718 (12,6)	848 (14,7)	0,001	0,84 (0,75-0,93)
		G	5.000 (87,4)	4.936 (85,3)		
<i>YTHDF3</i>	rs12115114	AA	73 (2,6)	70 (2,4)	0,78	1,02 (0,90-1,14)
		AG	711 (25,0)	715 (24,8)		
		GG	2.060 (72,4)	2.097 (72,8)		
		A	857 (15,1)	855 (14,8)	0,726	1,02 (0,92-1,13)
		G	4.831 (84,9)	4.909 (85,2)		
<i>ANKFN1</i>	rs17758761	CC	4 (0,1)	5 (0,2)	0,154	1,14 (0,95-1,38)
		AC	265 (9,3)	236 (8,2)		
		AA	2.589 (90,6)	2.648 (91,7)		
		C	273 (4,8)	246 (4,3)	0,181	1,13 (0,94-1,35)
		A	5.443 (95,2)	5.532 (95,7)		
<i>PTPRM</i>	rs4798571	AA	117 (4,1)	107 (3,7)	0,557	0,97 (0,87-1,08)
		AG	897 (31,5)	942 (32,6)		
		GG	1.835 (64,4)	1.838 (63,7)		
		A	1.131 (19,8)	1.156 (20,0)	0,818	0,99 (0,90-1,09)
		G	4.567 (80,2)	4.618 (80,0)		

CBLB	rs9657904	CC	107 (3,8)	143 (5,0)	0,053	0,90 (0,80-1,00)
		CT	866 (30,6)	912 (31,9)		
		TT	1.855 (65,6)	1.807 (63,1)		
		C	1.080 (19,1)	1.198 (20,9)	0,014	0,89 (0,81-0,98)
		T	4.576 (80,9)	4.526 (79,1)		
IL12B	rs6887695	CC	361 (12,6)	315 (10,9)	0,149	1,08 (0,97-1,20)
		CG	1.275 (44,6)	1.287 (44,5)		
		GG	1.223 (42,8)	1.293 (44,7)		
		C	1.997 (34,9)	1.917 (33,1)	0,039	1,08 (1,00-1,17)
		G	3.721 (65,1)	3.873 (66,9)		
IL12B	rs10045431	AA	303 (10,6)	337 (11,7)	0,024	0,89 (0,80-0,99)
		AC	1.199 (42,1)	1.266 (44,0)		
		CC	1.345 (47,2)	1.273 (44,3)		
		A	1.805 (31,7)	1.940 (33,7)	0,021	0,91 (0,84-0,99)
		C	3.889 (68,3)	3.812 (66,3)		
CNRI	rs1049353	AA	158 (5,8)	177 (6,5)	0,212	0,93 (0,84-1,04)
		AG	1.035 (37,8)	1.052 (38,7)		
		GG	1.546 (56,4)	1.488 (54,8)		
		A	1.351 (24,7)	1.406 (25,9)	0,145	0,94 (0,86-1,02)
		G	4.127 (75,3)	4.028 (74,1)		

IC: Intervalo de confianza; OR: *odds ratio*. * P calculada considerando el *carrier* del alelo minoritario o P alélica. Las ORs se han calculado teniendo en cuenta el efecto conferido por portar el alelo minoritario o el alelo minoritario.

La Tabla 15 muestra el análisis combinado de Mantel-Haenszel y el test de heterogeneidad entre las diferentes localizaciones. De los SNPs genotipados procedentes del primer GWAS ¹¹¹, el SNP rs1132200 localizado en el gen *TMEM39A* fue el único gen replicado en nuestro estudio ¹⁶⁰ realizado exclusivamente con población española [P = 0,001, OR (IC 95%) = 0,84 (0,75-0,93)] por lo que emergió como un gen de riesgo firmemente establecido, siendo el alelo común el que incrementa el riesgo a padecer esta enfermedad. El SNP rs1132200 causa un cambio aminoacídico de alanina a treonina en la posición 487 de la proteína resultante. Sin embargo, no se conoce casi nada acerca de *TMEM39A* y del papel biológico que podría desempeñar con respecto a la susceptibilidad a EM y se hacen necesarios estudios adicionales de mapeo fino y estudios funcionales. En un estudio de replicación multirracial a gran escala se identificó este gen como un loci de riesgo para lupus

eritematoso sistémico (LES, *systemic lupus erythematosus*)²⁰⁷, lo que aumenta el interés por este gen.

Tabla 15. Análisis combinado (*Mantel-Haenszel*) y heterogeneidad de los efectos del alelo minoritario de los marcadores previamente asociados con la susceptibilidad a EM, estratificados por localización.

Gen/SNP		Bilbao OR (IC 95%)	Madrid OR (IC 95%)	Andalucía OR (IC 95%)	Barcelona OR (IC 95%)	Análisis combinado P _{M-H} OR (IC 95%)	Análisis de heterogeneidad P (I ² %)
<i>PKN2</i> rs305217	T	0,83 (0,55-1,25)	0,71 (0,45-1,13)	0,9 (0,68-1,20)	1,51 (0,95-2,40)	0,64 0,94 (0,71-1,23)	0,13 (48)
<i>GTF2B</i> rs7538427	C	0,8 (0,53-1,20)	0,58 (0,36-0,94)	0,87 (0,66-1,16)	1,15 (0,70-1,89)	0,11 0,83 (0,66-1,04)	0,28 (22)
<i>EPHA4</i> rs1517440	C	1,19 (0,82-1,74)	0,83 (0,54-1,28)	1,07 (0,82-1,39)	1,16 (0,75-1,81)	0,5 1,06 (0,89-1,27)	0,63 (0)
<i>TMEM39A</i> rs1132200	A	0,93 (0,75-1,14)	0,78 (0,60-1,00)	0,81 (0,69-0,97)	0,081 (0,62-1,06)	0,001 0,84 (0,75-0,93)	0,70 (0)
<i>YTHDF3</i> rs12115114	A	1,04 (0,85-1,28)	0,93 (0,73-1,17)	1,09 (0,93-1,29)	0,91 (0,71-1,17)	0,76 1,02 (0,92-1,13)	0,54 (0)
<i>ANKFN1</i> rs17758761	C	0,94 (0,67-1,31)	1,6 (1,02-2,52)	1,07 (0,81-1,42)	1,37 (0,89-2,11)	0,17 1,16 (0,94-1,45)	0,23 (31)
<i>PTPRM</i> rs4798571	A	0,98 (0,82-1,17)	1,1 (0,89-1,09)	0,95 (0,82-1,09)	0,99 (0,79-1,24)	0,82 0,99 (0,90-1,08)	0,71 (0)
<i>CBLB</i> rs9657904	C	0,89 (0,74-1,06)	0,79 (0,63-0,98)	0,88 (0,76-1,02)	1,04 (0,82-1,31)	0,01 0,89 (0,81-0,97)	0,40 (0)
<i>IL12B</i> rs6887695	C	1,16 (0,99-1,35)	1,04 (0,87-1,25)	1,08 (0,96-1,22)	1,05 (0,86-1,28)	0,03 1,09 (1,01-1,17)	0,80 (0)
<i>IL12B</i> rs10045431	A	0,71 (0,61-0,82)	1,05 (0,88-1,27)	0,98 (0,86-1,11)	0,96 (0,79-1,17)	0,32 0,91 (0,76-1,09)	0,002 (80)*
<i>CNRI</i> rs1049353	A	0,93 (0,79-1,09)	0,94 (0,77-1,16)	1,05 (0,91-1,21)	0,72 (0,58-0,89)	0,23 0,91 (0,79-1,06)	0,04 (64)†

La tabla muestra el análisis combinado de *Mantel-Haenszel* y el test de heterogeneidad entre las diferentes colecciones. * Los resultados tras la exclusión de la muestra heterogénea de Bilbao son: P_{M-H} = 0,89, OR_{M-H} (IC 95%) = 0,99 (0,91-1,09). † Los resultados tras la exclusión de la muestra heterogénea de Barcelona son: P_{M-H} = 0,73, OR_{M-H} (IC 95%) = 0,98 (0,89-1,08), con una heterogeneidad final P (I²) de 0,77 (0) y 0,47 (0), respectivamente. IC: intervalo de confianza; OR: odds ratio.

En concordancia con los datos que aquí presentamos ¹⁶⁰, los SNPs rs4798571 y rs1517440 localizados respectivamente en los genes *PTPRM* y *EPHA4*, no alcanzaron una asociación significativa en la fase de replicación del primer GWAS ¹¹¹, aunque análisis combinados posteriores evidenciaron asociaciones marginales ²⁰⁶ que justificaban una confirmación adicional, ahora llevada a cabo, con lo que se descarta la asociación a EM de ambos SNPs. Otros dos SNPs estudiados, el rs10045431 y el rs1049353 localizados en los genes *IL12B* y *CNR1*, respectivamente, mostraron una heterogeneidad significativa debido a una de las poblaciones; sin embargo, una vez excluida del estudio correspondiente se observó una homogeneidad total en ambos casos ($I^2 = 0\%$), pero no se encontró una asociación significativa (Tabla 15).

El umbral de significación en los GWAS ($p < 10^{-7}$ - 10^{-8}) es equivalente al valor P nominal ($<0,05$) cuando se estudia un único marcador candidato y, por tanto, ambos tipos de estudio requieren validación de los descubrimientos originales. Sin embargo, las estrictas correcciones impuestas en un GWAS, a veces excluyen la identificación de verdaderas señales de asociación. Una de estas señales que no alcanzó la significación en previos GWAS en EM parece ser el gen *CBLB*, descrito como un marcador asociado en la población italiana ^{109, 110}. Sin embargo, la posibilidad de un efecto restringido a esa población alentaba su replicación. Nuestros datos demuestran que el alelo común del rs9657904 del gen *CBLB* se asocia significativamente con EM, pero con un efecto de menor magnitud (Tabla 15, OR = 1,12) que el previamente observado en la población italiana (OR = 1,4). El gen *CBLB*, que codifica una ubiquitin ligasa que regula los umbrales de activación de LTs y LBs, también se ha encontrado significativamente asociado con otras enfermedades inmunomediadas como la DMT1 en un estudio danés con 480 familias ²⁰⁸, aunque dicha asociación no se pudo confirmar más tarde en una cohorte mayor de pacientes ²⁰⁹.

Los dos SNPs estudiados del gen *IL12B* (rs6887695 y rs10045431), que codifica la subunidad p40 de las IL-12 e IL-23, muestran una asociación nominal con EM en este estudio [$P = 0,001$; OR (IC 95%) = 0,84 (0,75-0,93) y $P = 0,001$; OR (IC 95%) = 0,84 (0,75-0,93), respectivamente, Tabla 14]. Ambos SNPs se seleccionaron debido a que previamente se habían descrito efectos en otras enfermedades inmunomediadas en estudios realizados en poblaciones de ascendencia europea ^{101, 103, 104} y asiática ^{100, 102}. En psoriasis, se observó en diversos estudios independientes la asociación del rs6887695, aunque parece existir una discordancia sobre los alelos asociados. El alelo minoritario del SNP rs6887695 también se

asoció con un alto riesgo a colitis ulcerosa en población española [$P = 0,007$; OR (IC 95%) = 1,31 (1,07-1,60)]²¹⁰ y la implicación del gen *IL12B* en esta enfermedad se ha validado recientemente a través de meta-análisis de GWAS²¹¹. En un estudio de replicación a gran escala, el gen *IL12B* también destacó como un gen que podría contribuir al riesgo a LES²¹². El análisis combinado para este SNP realizado en nuestro estudio, muestra la predisposición conferida por el alelo minoritario (Tabla 15), tal como se había observado en los estudios publicados de otras enfermedades autoinmunes.

El otro SNP estudiado del gen *IL12B*, el rs10045431, había mostrado anteriormente protección a psoriasis²¹³ y espondilitis anquilosante¹⁰⁴, pero en nuestro estudio se observó una gran heterogeneidad entre las cohortes ($I^2 = 80$). Sin embargo, cuando corregimos esta heterogeneidad no pudimos detectar la asociación de este SNP con EM (Tabla 15). Estos datos concuerdan con los resultados de Begovich *et al.*²¹⁴, que tampoco encontraron una transmisión distorsionada en alelos de SNPs del gen *IL12B* en estudios familiares en EM, y con los publicados por Alloza *et al.*²¹⁵. Sin embargo, contrastan con los resultados publicados por Van Veen *et al.* que sí observaron asociación en su estudio realizado en población danesa²¹⁶, revelando por tanto un falso positivo.

Además, en el presente estudio nos propusimos estratificar a los individuos teniendo en cuenta el principal alelo de susceptibilidad a EM, el alelo *HLA-DRB1*15:01* y observamos que la comparación entre pacientes y controles *HLA-DRB1*15:01* positivos mostraba diferencias significativas en los SNPs estudiados de la *IL12B* [rs6887695: $P = 0,05$; OR (IC 95%) = 1,18 (0,99–1,41) y rs10045431: $P = 0,001$; OR (IC 95%) = 0,76 (0,64–0,9)]. Por otro lado, se había observado una asociación específica con hombres en un estudio en enfermedad de Graves²¹⁷ y la estratificación por género de nuestros datos también reveló un efecto predominante en hombres con EM [rs6887695: $P = 0,0075$, OR (IC 95%) = 1,21 (1,05–1,39) y rs10045431: $P = 0,035$; OR (IC 95%) = 0,86 (0,75–0,99)]. Los resultados relativos a *IL12B* en EM nos permiten establecer una conclusión firme, y en este punto resulta interesante recalcar que el gen que codifica la subunidad específica de la IL-12, el gen *IL12A*, se ha identificado también como un nuevo locus de susceptibilidad a EM¹⁷⁶.

Finalmente, nos propusimos estudiar el gen *CNRI*, que codifica un receptor con una amplia distribución en el SNC. Algunas publicaciones habían descrito que el gen *CNRI* influía en la progresión de la enfermedad en EMRR²¹⁸ y también se asoció con la predisposición a EMPP⁹⁸. La variante rs1049353 que conduce a una sustitución no sinónima

(Thr453Thr, 1359G/A) en el gen *CNR1* ha sido extensamente estudiada. Recientemente, se ha encontrado asociada con colitis ulcerosa y con enfermedad de Crohn ²¹⁹ y nos propusimos estudiar la asociación de este polimorfismo con EM en base a la genética compartida existente entre las diferentes enfermedades autoinmunes. Sin embargo, nuestros resultados no parecen apoyar un efecto de esta variante en la predisposición a la enfermedad [rs1049353: P = 0,23; OR (IC 95%) = 0,91 (0,79–1,06), Tabla 15].

Nuestros resultados no indican un papel significativo de los genes *PKN2*, *GTF2B*, *EPHA4*, *YTHDF3*, *ANKFN1*, *PTPRM* y *CNR1* en la susceptibilidad a EM, pero también refuerzan la necesidad de replicar señales sugerentes de asociación detectadas a través de los GWAS, ya que en ellos asociaciones espurias pueden alcanzar valores de significación respetables y solo se pueden distinguir después de un ejercicio de validación ^{114, 220}.

Mientras nuestro trabajo estaba bajo revisión, se publicó un destacado GWAS incluyendo 9.772 casos y 17.376 controles ¹²⁰. En él se identificaron 29 loci adicionales de susceptibilidad que se añadieron a los más de 20 loci de riesgo ya descritos. Nuestros resultados son completamente concordantes con este hallazgo ya que *TMEM39A*, *IL12B* y *CBLB* están listados entre los genes de susceptibilidad confirmados, pero no los restantes genes estudiados en el presente estudio ¹⁶⁰.

4. Estudio de cinco genes sugerentes de asociación tras GWAS

En el GWAS publicado por Sawcer *et al.*¹²⁰, 5 SNPs mostraron una diferencia estadísticamente significativa entre enfermos y controles aunque dicha significación no alcanzó los criterios de significación a nivel del genoma completo. Estos son: el rs228614 en *MANBA*, el rs630923 en *CXCR5*, el rs2744148 en *SOX8*, el rs2744148 en *RPS6KB* y el rs6062314 en *ZBTB46*. Dada la carencia de significación estadística, quisimos realizar una validación independiente para discernir el supuesto rol de estos loci en el riesgo a EM. Para este fin, se ha analizado la asociación de estos 5 SNPs con el riesgo a EM en un estudio multicéntrico incluyendo 8.499 enfermos, 8.765 controles no relacionados y 958 tríos (un enfermo y ambos progenitores) (Tabla 16). Este estudio tenía aproximadamente un 80% de potencia para detectar una OR de 1,10 para SNPs con MAFs de 0,13.

El meta-análisis de efectos fijos de las colecciones de validación mostró una asociación altamente significativa entre los 5 SNPs estudiados y el riesgo a EM: rs228614 [*MANBA*; $P = 2,4 \times 10^{-6}$; OR (IC 95%) = 0,91 (0,87-0,95)], rs630923 [*CXCR5*; $P = 1,2 \times 10^{-4}$; OR (IC 95%) = 0,89 (0,89-0,95)], rs2744148 [*SOX8*; $P = 1,8 \times 10^{-6}$; OR (IC 95%) = 1,14 (1,08-1,20)], rs180515 [*RPS6KB1*; $P = 5,2 \times 10^{-7}$; OR (IC 95%) = 1,12 (1,07-1,17)], y rs6062314 [*ZBTB46*; $P = 4,3 \times 10^{-3}$; OR (IC 95%) = 0,90 (0,83-0,97)]. El efecto observado fue similar al originalmente descrito¹²⁰ y no se encontraron evidencias sustanciales de heterogeneidad intraestudio para ninguno de los cinco SNPs estudiados (Tabla 16 y Figura 12).

Nuestro estudio¹⁶¹ muestra que variantes genéticas comunes ligadas a los genes *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* y *ZBTB46* están asociadas con el riesgo a sufrir EM y tras combinar nuestros datos con los originalmente descritos se alcanza una significación a nivel del genoma completo. Así, nuestros resultados proporcionan una fuerte evidencia de que estos loci representan factores de riesgo genéticos genuinos para EM.

Tabla 16. Resultados de la asociación de cinco loci y el riesgo a esclerosis múltiple estudiando 20.138 sujetos descendientes de europeos.

Gen	SNP	Colecciones de validación				Estudio original		P** Combinada
		P*	OR (IC 95%)	I ² (IC 95%)	P ₀	P**	OR	
MANBA	rs228614 (A/ <u>G</u>)	2,4x10 ⁻⁶	0,91 (0,87–0,95)	37 (0–71)	0,12	1,4x10 ⁻⁷	0,92	3,4x10 ⁻¹²
CXCR5	rs630923 (A/ <u>C</u>)	1,2x10 ⁻⁴	0,89 (0,84–0,95)	1 (0–65)	0,425	2,8x10 ⁻⁷	0,89	4,7x10 ⁻¹⁰
SOX8	rs2744148 (<u>G</u> /A)	1,8x10 ⁻⁶	1,14 (1,08–1,20)	0 (0–14)	0,915	8,4x10 ⁻⁸	1,12	1,6x10 ⁻¹²
RPS6KB1	rs180515 (G/ <u>A</u>)	5,2x10 ⁻⁷	1,12 (1,07–1,17)	28 (0–66)	0,198	8,8x10 ⁻⁸	1,09	2,3x10 ⁻¹³
ZBTB46	rs6062314 (C/ <u>T</u>)	4,3x10 ⁻³	0,90 (0,83–0,97)	31 (0–68)	0,169	1,3x10 ⁻⁷	0,86	2,3x10 ⁻⁸

Los meta-análisis de efectos fijos de los SNPs estudiados en las colecciones de validación se realizaron en PLINK. El umbral de significación nominal se estableció aplicando la *corrección de Bonferroni* para cinco tests en $P < 0,01$. Todos los valores de P son de una cola con respecto a la dirección esperada del efecto basada en el estudio original ¹²⁰. Los resultados de la asociación del estudio original ¹²⁰ y el estudio actual se combinaron usando METAL ¹⁷⁵. Los nombres de los alelos se muestran con alelo minoritario/alelo mayoritario basado en las frecuencias en el conjunto de datos de validación total y los alelos de riesgo están subrayados. Los paréntesis que siguen al nombre del gen indican la localización del SNP respecto al gen, las pares de bases (pb) indican la distancia *upstream* (5') o *downstream* (3') al transcrito primario. IC: intervalo de confianza; OR: *odds ratio*; UTR: *untranslated region*; *: una cola; **: dos colas.

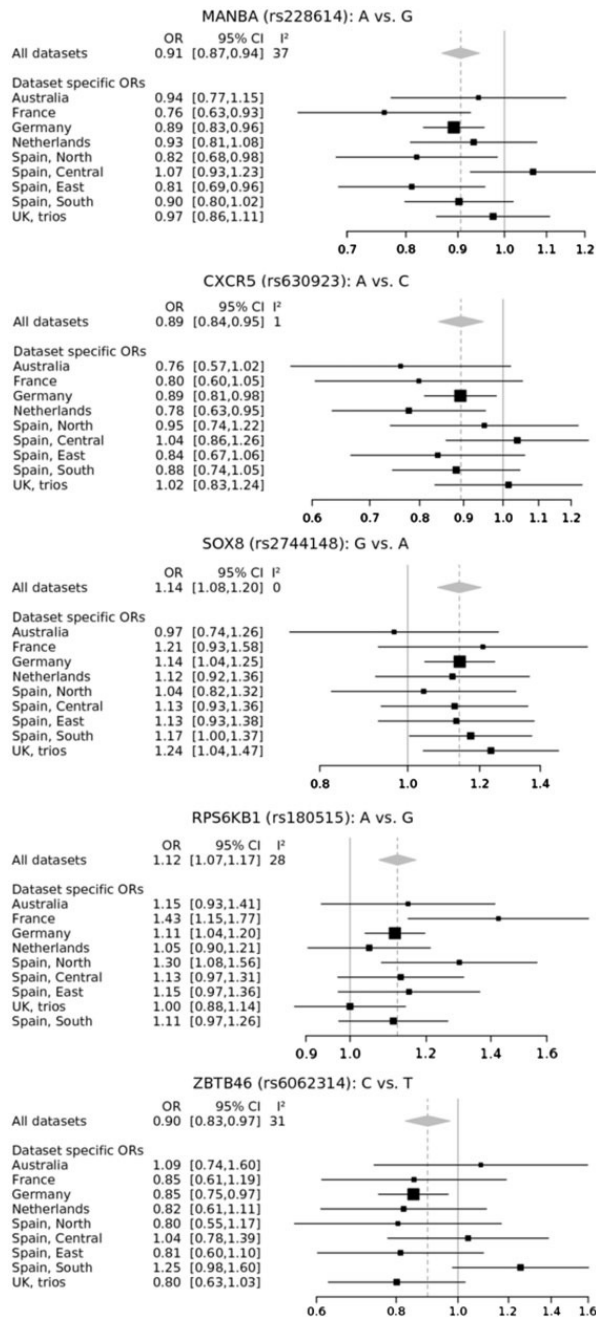


Figura 12. Meta-análisis de las colecciones de datos estudiando la asociación entre los loci MANBA, CXCR5, SOX8, RPS6KB1 y ZBTB46 y el riesgo a padecer EM en poblaciones descendientes de europeos. El eje de las x indica las *odds ratio* (OR). Las ORs específicas de estudio (cuadrados negros) y el 95% de intervalo de confianza (CI, líneas) se calcularon usando un modelo aditivo (rombos grises) en base a un modelo de efectos fijos. Los *forest plot* se generaron usando una versión personalizada del paquete estadístico *rmeta* en lenguaje R²²¹.

Como en el caso de la mayoría de las asociaciones genéticas descritas hasta el momento, los mecanismos moleculares subyacentes a estos resultados aún no se han estudiado. Esto implica que se necesitan estudios adicionales para clarificar si los SNPs aquí estudiados están directamente involucrados en alterar la expresión génica, la función de la proteína o si dichos efectos son ejercidos por otras variantes correlacionadas. Por ejemplo, el SNP rs180515 en el extremo 3' UTR de *RPS6KB1* se localiza en la región semilla (*seed*) de un sitio de unión previsto para los microRNAs (miRNAs) hsa-miR-3616-5p y hsa-miR-573, en los transcritos codificantes de proteínas ENST00000443572 y ENT0000039302 del *RPS6KB1*. La presencia del alelo minoritario G en ambos transcritos mRNA puede afectar la unión de ambos miRNAs debido a una alteración en el apareamiento normal y la aparición de un apareamiento tipo *wobble*. Los miRNAs son RNAs de una sola hebra de unos 20 a 24 nucleótidos de longitud que regulan la expresión de mRNAs dianas a nivel transcripcional. En este caso, se supone que esto puede disminuir la afinidad de la unión del hsa-miR-3616-5p/hsa-miR-573 al 3'UTR del *RPS6KB1* y puede, por tanto, incrementar la transcripción de *RPS6KB1*. Distintos trabajos ²²²⁻²²⁵ han descrito la forma en la que estos miRNAs alteran la regulación del sistema inmune en EM y podrían contribuir a mejorar la comprensión de los mecanismos patológicos que conducen a esta enfermedad.

El SNP intrónico rs228614 en *MANBA* está en LD con 2 SNPs no sinónimos en el mismo gen [rs2866413 (p.Thr701Met), $r^2 = 0,87$ y rs227368 (p.Val253Leu), $r^2 = 0,74$, basado en datos para población CEU del "Proyecto 1.000 Genomas" ²²⁶] que pueden afectar la función de la proteína. Sin embargo, y posiblemente mucho más importante, es el hecho de que de los 5 loci estudiados, *MANBA* es el único que contiene SNPs (incluyendo el rs228614) mostrando fuertes efectos *cis* en la expresión del RNA mensajero, de acuerdo con datos recientemente publicados ²²⁷. El SNP intrónico rs6062314 en *ZBTB46* solo muestra un LD reducido con variantes potencialmente funcionales, es decir, con SNP no sinónimos, en *LIME1* [*Lck interacting transmembrane adaptor 1*, SNP rs1151625 (p.Pro211Leu), $r^2 = 0,39$] y en *ZGPAT* [*zinc finger, CCCH-type with G patch domain*, SNP rs1291212 (p.Ser61Arg), $r^2 = 0,31$]. Además, *TNFRSF6B* (*tumour necrosis factor receptor superfamily, member 6b, decoy*) se localiza también en esta región cromosómica y representaría un mejor candidato en base a sus implicaciones en la función de los LTs ²²⁸.

Finalmente, el rs630923 se localiza en un sitio de unión al factor de transcripción *CXCR5*, un factor nuclear potenciador de las cadenas ligeras kappa de los LBs en una región

hipersensible para DNasa I²²⁹. Se ha previsto que el rs630923 altera la unión del NFκB²³⁰ y podría así, afectar potencialmente la transcripción de *CXCR5*.

Es conveniente enfatizar que las consecuencias funcionales potenciales mencionadas anteriormente se basan en estudios *in silico* y requieren estudios experimentales y de validación. Además, es posible también que variantes de DNA raras, desconocidas hasta ahora, subyagan o contribuyan a las señales de asociación observadas.

5. Mapeo fino y análisis funcional del gen de riesgo *CD6*

En este estudio ¹⁶⁴ nos propusimos identificar las variantes genéticas del *CD6* asociadas a EM a través de un ejercicio de mapeo fino. Se llevó a cabo un cribado de SNPs del *CD6* en dos fases para identificar la supuesta variante causal o los haplotipos que contribuyen a la mayor susceptibilidad a EM. Se analizaron los efectos de los haplotipos asociados a la expresión de *CD6* en la superficie celular de LTs y células NK por citometría de flujo y estudiamos las diferencias en proliferación y en producción de citoquinas (IFN- γ e IL-17) de acuerdo con su haplotipo.

5.1. Estudio de variantes de riesgo

La primera etapa del estudio se llevó a cabo en una colección de individuos del País Vasco en un total de 823 pacientes con EM y 814 controles sanos. Se seleccionaron 9 SNPs marcadores de haplotipo (*Hapmap release # 27*) para cubrir la variabilidad del gen *CD6*. También se incluyeron los SNPs no sinónimos rs11230563 (R225W) y rs2074225 (A257V) en el segundo dominio SRCR (exón 4); y el rs12360861 (A271T) en el tercer dominio SRCR (exón 5). Finalmente, se incluyeron los SNPs rs17824933 y el rs650258. Por un lado, el rs17824933 se seleccionó porque había mostrado el riesgo más alto (OR = 1,34) en un estudio de validación ²³¹ de las señales de asociación identificadas por De Jager *et al.* ¹¹⁶. Por otro, el SNP rs650258 se incluyó en base a la asociación observada en un GWAS de EM ¹²⁰.

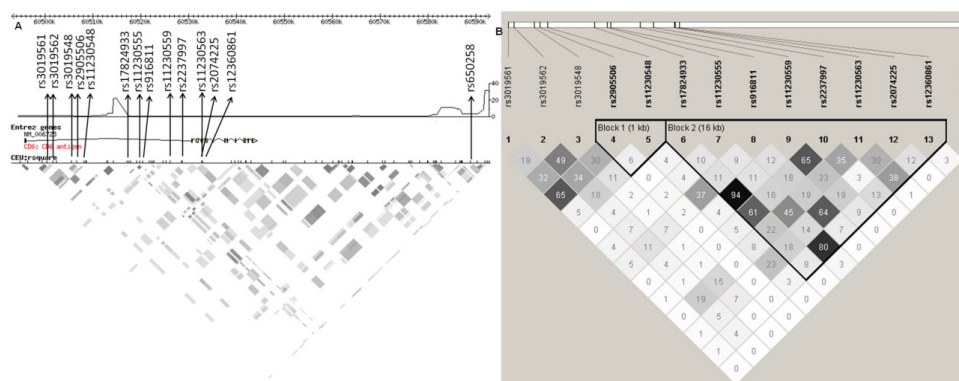


Figura 13. Localización de los 14 SNPs estudiados y sus patrones de disequilibrio de ligamiento en el locus del *CD6*. (A) Imágenes de bloques de LD con las localizaciones aproximadas de los 14 SNPs marcadores elegidos para este estudio, puntos de recombinación y patrones de LD dentro del *plot* de LD para población CEU (*Hapmap relase # 27*) del gen *CD6*, y (B) patrones de LD en la población del País Vasco usando el método del intervalo de confianza (generado usando el software *Haploview*²³²).

La potencia de nuestro estudio fue superior al 92% para cada SNP (salvo para el rs12630861 con un 72%) con una OR = 1,34, similar a la OR del rs17824933 en el estudio de Swaminathan *et al.*²³¹ y para las frecuencias de población CEU. Todos los SNPs estudiados estaban en HWE salvo el rs2237997 (P = 0,015) (Tabla 17).

Tabla 17. Análisis de asociación de 14 SNPs del *CD6* en el País Vasco.

SNP ^a	Localización	Cambio aminoacídico	Frecuencias ^b	P	OR (IC 95%)
			EM/Controles		
rs3019561 (<u>T/G</u>)	Intrón 1	-	0,25/0,24	0,44	1,07(0,91–1,25)
rs3019562 (<u>C/G</u>)	Intrón 1	-	0,51/0,47	0,04	1,16(1,01–1,33)
rs3019548 (<u>G/A</u>)	Intrón 1	-	0,40/0,38	0,16	1,11(0,96–1,28)
rs2905506 (<u>T/C</u>)	Intrón 1	-	0,26/0,25	0,36	1,08(0,92–1,26)
rs11230548 (<u>A/C</u>)	Intrón 1	-	0,86/0,82	0,008	1,28(1,06–1,56)
rs17824933 (<u>G/C</u>)	Intrón 1	-	0,29/0,25	0,007	1,24(1,06–1,45)
rs11230555 (<u>A/C</u>)	Intrón 1	-	0,78/0,77	0,57	1,05(0,89–1,25)
rs916811 (<u>A/G</u>)	Intrón 1	-	0,79/0,72	3,3x10 ⁻⁵	1,40(1,21–1,63)
rs11230559 (<u>C/T</u>)	Intrón 1	-	0,30/0,25	0,004	1,25(1,07–1,46)
rs2237997 (<u>C/T</u>)	Intrón 1	-	0,38/0,35	0,07	1,14(0,99–1,32)
rs11230563 (<u>T/C</u>)	Exón 4	R225W	0,40/0,36	0,08	1,13(0,98–1,31)
rs2074225 (<u>C/T</u>)	Exón 4	A257V	0,71/0,63	3,1x10 ⁻⁶	1,40(1,21–1,63)
rs12360861 (<u>A/G</u>)	Exón 5	A271T	0,81/0,81	0,79	1,03(0,85–1,21)
rs650258 ^c (<u>T/C</u>)	3' Intergénico		0,56/0,51	0,0052	1,23(1,06–1,41)

^a Los alelos de riesgo están subrayados. ^b Valores correspondientes al alelo de riesgo.

Con los seis SNPs que mostraron una mayor significación se llevó a cabo una fase de validación en un total de 2.265 pacientes con EM y 2.600 controles de 3 colecciones diferentes. También se incluyó en esta fase el SNP no sinónimo rs11230563 (R225W), ya que se observó una tendencia hacia la asociación en el cribado inicial [P = 0,08; OR (IC 95%) = 1,13 (0,98-1,31)]. Tres SNPs emergieron con significación nominal del grupo de replicación: el rs17824933, el rs11230559 y el rs2074225 (Tabla 18). Además, los análisis combinados de todos los datos mostraron que el SNP no sinónimo rs650258 también se encontraba asociado.

Tabla 18. Análisis de replicación y combinado de los SNPs más significativos.

SNPs ^a	Replicación			Estudio combinado		
	P _{CMH}	P _{BD}	OR ^b (IC 95%)	P _{CMH}	P _{BD}	OR ^b (IC 95%)
rs11230548 (A/C)	0,64	0,25	0,97 (0,85–1,1)	0,85	0,17	1,01 (0,90–1,13)
rs17824933 (G/C)	0,02	0,44	1,14 (1,02–1,27)	0,005	0,58	1,16 (1,04–1,28)
rs916811 (A/G)	1,00	0,49	1,00 (0,89–1,13)	0,45	0,32	1,04 (0,94–1,16)
rs11230559 (C/T)	0,02	0,20	1,14 (1,02–1,27)	0,005	0,31	1,15 (1,05–1,27)
rs11230563 (T/C) ^b	0,77	0,69	0,98 (0,89–1,09)	0,88	0,83	0,99 (0,91–1,09)
rs2074225 (C/T)	0,03	0,53	1,14 (1,01–1,25)	0,0065	0,67	1,14 (1,04–1,25)
rs650258 (T/C)	0,08	0,46	1,085 (0,99–1,19)	0,003	0,30	1,12 (1,04–1,21)

CMH: test de *Cochran-Mantel-Haenszel*, BD: test de *Breslow-Day*, OR= *odds ratio*, IC= intervalo de confianza. ^a Los alelos de riesgo (con respecto al conjunto de datos del País Vasco) están subrayados.

^b rs11230563 se incluye porque es un SNP no sinónimo (R225W), y muestra una tendencia hacia la asociación en el primer cribado.

El análisis de los patrones de LD mostró que el rs11230559 está en fuerte LD con el rs17824933 en la colección del País Vasco ($D' = 0,99$; $r^2 = 0,93$; Figura 14) y también en el conjunto de datos combinado (no mostrado), de acuerdo con los datos del “Proyecto 1.000 Genomas” ($D' = 1$; $r^2 = 0,93$; Figura 14). Además, estos dos SNPs están en fuerte LD con dos SNPs no sinónimos: el rs11230562 (T217M) en el dominio 2 del SRCR (exón 4) ($D' = 1$; $r^2 = 0,94$ con el rs17824933 y $D' = 1$; $r^2 = 0,8$ con rs11230559) y el rs2074233 (G606S) en el dominio citoplásmico (exón 11) ($D' = 0,94$; $r^2 = 0,84$ con el rs17824933 y $D' = 1$; $r^2 = 0,898$ con el rs11230559). La asociación de estos SNPs era más fuerte en la colección de País Vasco en comparación con los grupos de replicación, lo que podría estar relacionado con factores tales como la heterogeneidad de la enfermedad y una menor heterogeneidad genética en el País Vasco. Por otro lado, el rs650258 apenas mostró LD con alguno de los 3 SNPs no sinónimos: el rs11230563 ($D' = 0,224$; $r^2 = 0,038$), rs2074225 ($D' = 0,095$; $r^2 = 0,005$) y rs12360861 ($D' = 1$; $r^2 = 0,149$) (<http://www.broadinstitute.org/mpg/snap/ldsearch.php>)²³³.

SNP	rs3019561	rs3019562	rs3019548	rs2905506	rs11230548	rs17824933	rs11230555	rs916811	rs11230559	rs2237997	rs11230563	rs2074225	rs12360861	rs650258
rs3019561	1													
rs3019562	0.61 / 0.24 0.76 / 0.19	1												
rs3019548	0.64 / 0.41 0.79 / 0.32	0.95 / 0.59 0.87 / 0.48	1											
rs2905506	0.89 / 0.571 0.82 / 0.63	1 / 0.48 0.99 / 0.34	0.83 / 0.50 0.74 / 0.30	1										
rs11230548	1 / 0.20 0.91 / 0.04	1 / 0.3 0.98 / 0.18	1 / 0.20 0.97 / 0.11	1 / 0.14 1 / 0.06	1									
rs17824933	0.08 / 0 0.04 / 0	0.23 / 0.006 0.35 / 0.03	0.08 / 0 0.29 / 0.01	0.02 / 0 0.05 / 0	0.42 / 0.02 0.81 / 0.05	1								
rs11230555	0.18 / 0.01 0.22 / 0.04	0.28 / 0.02 0.48 / 0.06	0.04 / 0 0.16 / 0.01	0.02 / 0 0.17 / 0.01	1 / 0.07 0.94 / 0.04	1 / 0.03 0.99 / 0.10	1							
rs916811	0.72 / 0.11 0.7 / 0.05	0.61 / 0.12 0.59 / 0.12	0.62 / 0.08 0.62 / 0.09	0.36 / 0.02 0.49 / 0.02	0.58 / 0.30 0.79 / 0.40	1 / 0.123 1 / 0.11	1 / 0.076 1 / 0.09	1						
rs11230559	0.25 / 0.006 0.03 / 0	0.31 / 0.01 0.34 / 0.03	0.25 / 0.006 0.27 / 0.01	0.23 / 0.004 0.08 / 0	0.19 / 0.005 0.74 / 0.04	1 / 0.85 0.99 / 0.93	1 / 0.033 1 / 0.11	1 / 0.15 1 / 0.12	1					
rs2237997	0.14 / 0.01 0.13 / 0	0.11 / 0.004 0.17 / 0.01	0.20 / 0.009 0.23 / 0.01	0.09 / 0.003 0.11 / 0	0.29 / 0.03 0.72 / 0.06	1 / 0.35 0.98 / 0.60	1 / 0.079 1 / 0.16	0.92 / 0.29 0.97 / 0.17	1 / 0.42 0.99 / 0.65	1				
rs11230563	0.21 / 0.04 0.11 / 0	0.60 / 0.23 0.54 / 0.18	0.42 / 0.17 0.39 / 0.14	0.07 / 0.003 0.05 / 0	0.86 / 0.14 0.82 / 0.08	1 / 0.08 0.98 / 0.22	1 / 0.36 0.97 / 0.44	1 / 0.21 0.98 / 0.18	1 / 0.09 0.99 / 0.24	1 / 0.22 0.99 / 0.36	1			
rs2074225	0.03 / 0 0.27 / 0.01	0.29 / 0.05 0.32 / 0.06	0.32 / 0.04 0.47 / 0.07	0.12 / 0.004 0 / 0	0.46 / 0.10 0.79 / 0.25	1 / 0.23 0.99 / 0.17	0.76 / 0.08 0.98 / 0.13	0.94 / 0.47 0.99 / 0.63	0.89 / 0.22 0.99 / 0.18	0.306 / 0.06 0.33 / 0.03	0.93 / 0.34 0.99 / 0.30	1		
rs12360861	0.06 / 0.001 0.14 / 0.01	0.42 / 0.022 0.41 / 0.02	0.36 / 0.003 0.05 / 0	1 / 0.02 0.09 / 0	1 / 0.037 0.92 / 0.03	1 / 0.02 0.98 / 0.08	1 / 0.54 0.99 / 0.83	1 / 0.04 1 / 0.07	1 / 0.02 1 / 0.09	1 / 0.043 1 / 0.14	1 / 0.20 0.99 / 0.39	1 / 0.08 1 / 0.12	1	
rs650258	0.09 / 0.006 0.05 / 0	0.21 / 0.036 0.01 / 0	0.08 / 0.001 0.04 / 0	0.18 / 0.02 0 / 0	0.12 / 0.003 0.1 / 0	0.63 / 0.041 0.15 / 0	0.86 / 0.204 0.31 / 0.03	0.495 / 0.07 0.04 / 0	0.64 / 0.05 0.16 / 0	0.11 / 0.01 0.1 / 0	0.22 / 0.04 0.11 / 0.01	0.095 / 0.005 0 / 0	1 / 0.15 0.33 / 0	1

Figura 14. Valores de LD (D'/r^2) entre cada uno de los 14 SNPs genotipados en la población del País Vasco. Los valores en blanco son valores de D'/r^2 extraídos del “Proyecto 1.000 Genomas” en población CEU, mientras que los azules son los generados a partir de los datos del País Vasco usando el software *Haploview*.

Los análisis de haplotipos se realizaron en todas las colecciones. Para ello, se llevó a cabo un análisis de ventana deslizante usando dos, tres o cuatro marcadores en el que se realizan múltiples correcciones basadas en el número de SNPs analizados teniendo en cuenta su estructura de correlación. Además, se realizó un análisis de dos o tres marcadores usando al menos uno de los dos SNPs no sinónimos. El análisis de ventana deslizante mostró una fuerte asociación de combinaciones incluyendo los SNPs exónicos no sinónimos rs11230563 (R225W) y rs2074225 (A257V) ($P_{\text{corregida por 10.000 permutaciones}} = 1 \times 10^{-4}$) (Figura 15).

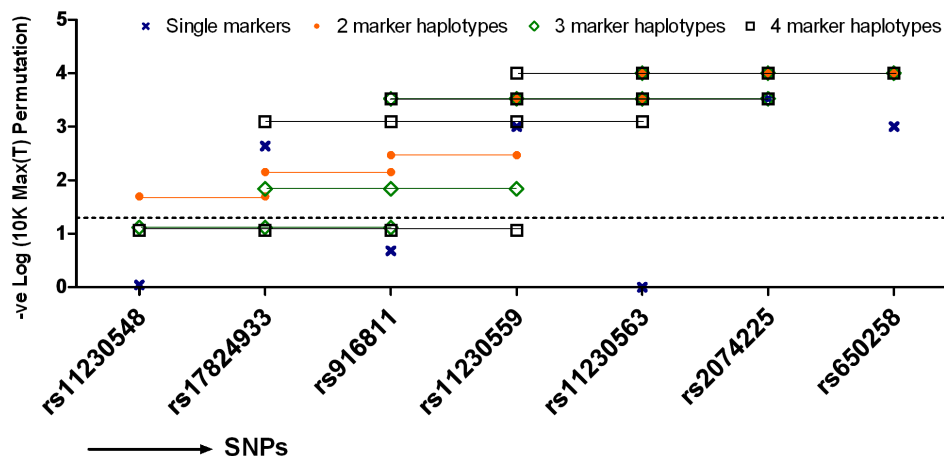


Figura 15. Análisis de ventana deslizante de los haplotipos usando dos o tres marcadores por medio de un análisis de 10.000 permutaciones. La asociación total de los dos/tres marcadores adyacentes se estudió mediante un análisis de ventana deslizante cambiando un marcador cada vez. El análisis se realizó usando una combinación de las tres colecciones de replicación y la colección de datos original con un total de 3.008 enfermos con EM y 3.414 controles para los seis SNPs seleccionados para la fase de replicación. Los valores de $-\log^{10}$ de P para la asociación de cada SNP con EM se muestran como cruces azules sin conectar, y las asociaciones de haplotipos basadas en ventana deslizante de dos, tres y cuatro SNPs se representan como líneas naranjas, verdes o negras, respectivamente.

La adición de los SNPs rs17824933 o rs1230559 a los marcadores haplotípicos rs11230563-rs2074225 no modificó las frecuencias haplotípicas significativamente ni incrementó la fuerza de asociación (datos no mostrados), indicando que la asociación era explicada en esencia por el haplotipo rs11230563-rs2074225 ($P_{\text{corregida por 10.000 permutaciones}} = 1 \times 10^{-4}$) (Figura 15). El uso de un *tagging* agresivo en haplotipos de dos o tres marcadores mostró que el rs11230563 era equivalente a los marcadores SNPs rs916811 y/o rs11230559 junto con el rs2074225 ($r^2 = 0,95$) y también por el rs916811 y/o el rs17824933 con el rs2074225 ($r^2 = 0,89$). De forma semejante, el rs2074225 era equivalente a los marcadores rs11230563 y rs11230559/rs17824933 ($r^2 = 0,93/0,88$; respectivamente). Esto indicaba que los dos SNPs no sinónimos eran equivalentes a los SNPs intrónicos en combinación con cualquiera de los SNPs no sinónimos.

Posteriormente se llevó a cabo un análisis de regresión logística asumiendo efectos aditivos de dosis alélicas para estudiar los efectos independientes. El análisis reveló efectos

independientes ejercidos por los SNPs rs2074225 y rs650258, y fuertes efectos aditivos con una $P = 6,2 \times 10^{-7}$ condicionando rs11230563 a rs2074225.

En resumen, identificamos dos nuevos SNPs asociados con el riesgo a EM, que son el rs11230559 localizado en la región intrónica y el rs2074225, un SNP no sinónimo localizado en el segundo dominio SRCR y validamos la asociación del rs650258¹²⁰ con la susceptibilidad a EM. Por lo tanto, el análisis de haplotipos por un análisis de ventana deslizante mostró una fuerte asociación de los dos SNPs no sinónimos del dominio 2 (exón 4) que se confirmaron por los análisis de regresión logística.

5.2. Determinación de CD6 en función del haplotipo rs11230563-rs2074225

La proteína CD6 es un miembro del grupo B de la superfamilia de receptores *scavenger* ricos en cisteína (SRCR-SF, *scavenger receptor cysteine-rich super family*)²³⁴ encontrado en timocitos, LTs maduros, algunos LBs y subgrupos de células NK, aunque también se expresa en ciertas partes del cerebro como el cerebelo, los ganglios basales, el tálamo, el cuerpo amigdaloides y los lóbulos cerebrales²³⁵⁻²³⁸. Por lo tanto, también nos propusimos analizar los cambios de expresión del CD6 en diferentes tipos celulares del sistema inmunitario de acuerdo a los genotipos de los SNPs no sinónimos rs11230563 y rs2074225 marcadores del *CD6* de forma individual o formando haplotipos. Los análisis de FACS se realizaron con linfocitos de pacientes con EM genotipados para estos dos SNPs. Nuestro análisis incluyó además de los LTs CD4⁺ y CD8⁺ estudiados por Kofler *et al.*²³⁹, a las células NK, células NKT y subgrupos de LTs efectores en un intento de identificar el tipo celular más fuertemente influenciado por los genotipos de SNPs no sinónimos/haplotipos del *CD6*.

Nuestros resultados mostraron que CD6 se expresaba a distintos niveles en los subgrupos de LTs CD4⁺, LTs CD8⁺ y células NKT (Figura 16). Estas diferencias se observaron entre los LTs CD4⁺ ($P_{ANOVA} = 0,0008$) y CD8⁺ ($P_{ANOVA} = 0,001$), pero también aunque con menor significación en células NKT ($P_{ANOVA} = 0,02$). En estos tres subgrupos de linfocitos, los niveles de expresión de CD6 varían significativamente con el haplotipo protector CC (rs11230563-rs2074225) expresando niveles más altos de CD6 que el haplotipo CT de riesgo y el haplotipo TT que proporciona un riesgo leve (Figura 16). En los análisis se observó que el haplotipo ancestral CC conducía a una mayor expresión de CD6 en superficie

comparado con los haplotipos de riesgo (CT y TT), en línea con las tendencias observadas por Kofler *et al.* en la isoforma de longitud completa ²³⁹ y el análisis de Heap *et al.* ²⁴⁰ que señalaba un desequilibrio alélico de 8 SNPs en la región del CD6 incluyendo el rs2074225, el rs11230562 y el rs17824933, inferido a través de patrones de LD. También observamos que la expresión de CD6 de las células con el haplotipo TT (riesgo leve) era más baja que aquella del haplotipo CT (de riesgo). Aunque el haplotipo CT es intermedio en términos de expresión de CD6 entre el haplotipo protector (CC) y leve riesgo (TT) en los LTs CD4⁺ y CD8⁺, su expresión muestra una tendencia más alta en LTs de memoria CD8⁺ y CD4⁺, aunque esa diferencia no llega a ser significativa. Así, nuestros datos sugieren que el riesgo genético conferido por CD6 es complejo y podría resultar en cambios en los niveles de expresión y unión del ligando o transducción de la señal vía los 2 SNPs no sinónimos rs11230563 (R225W) y rs2074225 (A257V) del dominio 2 del SRCR, además del rs2074233 (G606S) del dominio citoplásmico.

A nivel transcripcional, además de la forma de longitud completa del CD6, se han publicado un total de seis isoformas que diversifican los dominios citoplásmicos ^{241, 242}. El ligando de CD6 es la molécula de adhesión celular del leucocito activado (ALCAM, *activated leukocyte cell adhesion molecule*), encontrado en el epitelio tímico y en la capa epitelial de la BHE. ALCAM se une al tercer dominio SRCR del CD6 y se ha observado que esta interacción facilita la trans migración de LTs CD4⁺ a través de la BHE ²⁴³. Además, se ha descrito que CD6 está implicado en la maduración de la sinapsis inmunológica por asociación a la región del complejo de activación supramolecular central ²⁴¹. Sin embargo, la isoforma CD6ΔD3 carente del dominio de unión a ALCAM, que estaba sobrerregulada en la activación de LTs, no estaba localizada en la sinapsis inmunológica ^{241, 244}. Los estudios usando linfocitos muestran una mayor expresión de CD6 en timocitos en estado maduro (solo positivo) comparados con el estado inmaduro (doble negativo, doble positivo) y muestran la influencia negativa de CD6 en la tasa de apoptosis, sugiriendo su papel en la activación tímica ²³⁵. También se ha observado que la expresión del CD6ΔD3 era más alta en los timocitos maduros comparados con los inmaduros ²⁴¹. Sin embargo, en nuestro estudio debido a que los anticuerpos monoclonales utilizados en este estudio por FACS reconocían el dominio 1 de la región del SRCR, no fue posible la discriminación entre las distintas isoformas de CD6.

Se debería analizar la contribución de la isoforma carente del dominio de unión al ligando (CD6ΔD3), cuya expresión está inversamente correlacionada con aquella de la forma de longitud completa²³⁹ para dilucidar la función del CD6. Kofler *et al.*²³⁹ observaron diferencias no significativas en la expresión del CD6 de la forma de longitud completa entre los genotipos, pero encontraron una alta expresión de CD6ΔD3 en los LTs CD4⁺ y CD8⁺ en individuos homocigotos para el alelo de riesgo rs17824933GG. En este análisis de correlación de los genotipos del rs17824933 con la expresión de dos isoformas del dominio extracelular del CD6 (longitud completa y CD6ΔD3) en LTs CD4⁺ y CD8⁺ de controles sanos no se encontraron diferencias significativas. Sin embargo, la comparación de la expresión relativa de las dos isoformas mostró que el alelo de riesgo (rs17824933G) se asociaba con una expresión más alta de la isoforma CD6ΔD3 que carecía del dominio de unión al ligando²³⁹.

En un análisis más profundo, las mayores diferencias se encontraron en el subgrupo *näive* (células *näive* CD4⁺, $P_{ANOVA} = 0,0001$; células *näive* CD8⁺, $P_{ANOVA} = 0,0001$) cuando se separaron los subgrupos de LTs CD4⁺ y CD8⁺ de acuerdo a sus fenotipos *näive*, de memoria o efectoras. El subgrupo de LTs CD4⁺ de memoria también mostró diferencias significativas ($P_{ANOVA} = 0,01$), pero no se encontraron diferencias en los subgrupos efectoras y en las células efectoras diferenciadas terminalmente (CD27⁻ CD28⁻). En nuestro estudio, el haplotipo estudiado de CD6 parece tener un mayor efecto en las poblaciones de LTs CD4⁺ y CD8⁺ vírgenes que en células de memoria y efectoras centrales. Esto es importante porque CD6 tiene una elevada expresión en células tímicas y se piensa que ayuda a la selección positiva y evita la apoptosis²³⁵. Además, Singer *et al.*²³⁵ y Castro *et al.*²⁴¹ han mostrado que las dos isoformas presentaban alta expresión en células maduras. En el timo, CD6ΔD3 está infraexpresado en las células doble positivas, mientras que la expresión de la forma de longitud completa está favorecida, lo que contribuye a su supervivencia.

Las células NKT también expresaron CD6 y en esta población también se observaron las diferencias relacionadas con el haplotipo ($P_{ANOVA} = 0,02$) (Figura 16). Por otro lado, las células NK expresaron niveles mucho más bajos de CD6 en su superficie y no se observaron cambios significativos en relación a los haplotipos del *CD6*.

En resumen, nuestros datos muestran un efecto más pronunciado del haplotipo rs11230563-rs2074225 en comparación con los SNPs no sinónimos individuales del *CD6*.

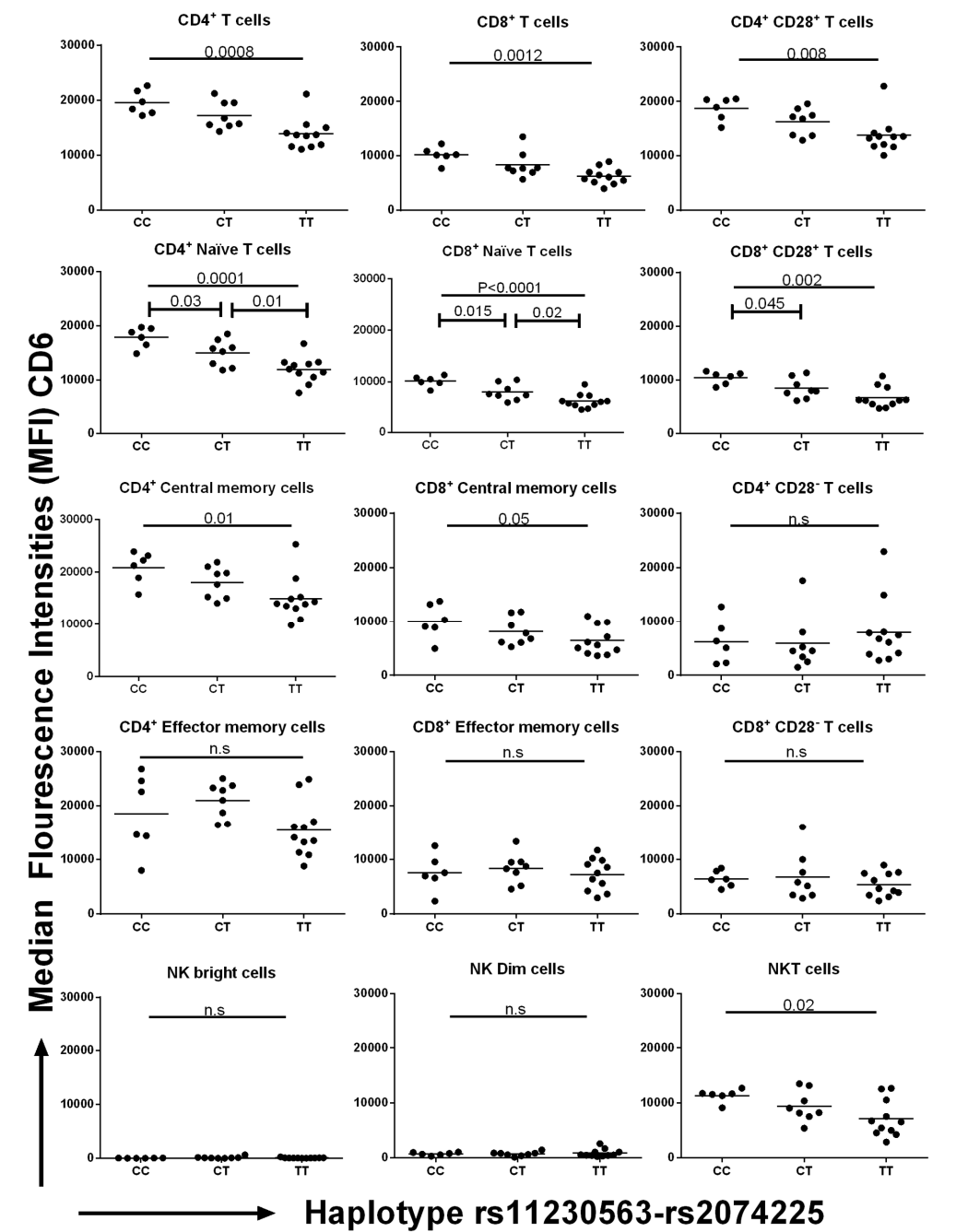


Figura 16. Expresión del CD6 en diferentes tipos celulares dependiendo del haplotipo rs11230563-rs2074225 presente. La expresión de CD6 en diferentes grupos celulares se analizó en un total de 27 muestras de PBMCs que representaban a cada uno de tres haplotipos de los dos SNPs no sinónimos: rs11230563 y rs2074225. El eje de ordenadas representa la intensidad media de fluorescencia (MFI) mientras que el eje de abscisas representa los tres diferentes haplotipos; es decir, el haplotipo CC que confiere protección, el CT asociado con riesgo y TT que confiere un riesgo leve. Los marcadores CD3 y CD56 se usaron para identificar células T ($CD4^+/CD8^+ CD3^+CD56^-$), células NK ($CD3^+ CD56^+$), y el grupo de células NKT ($CD3^+ CD56^+$); CD45RA, CD27 y CD28 se utilizaron para distinguir células vírgenes ($CD45RA^+ CD27^+ CD28^+$), células de memoria central ($CD45RA^- CD27^+$) y células de memoria efectoras ($CD45RA^- CD27^- CD28^+$), así como las células efectoras de memoria diferenciadas terminalmente (TEMRA, $CD45RA^+ CD27^- CD28^-$); mientras que CD56 y CD16 se utilizaron para identificar las células NK $CD56^{dim}$ ($CD56^{int} CD16^+$) y las $CD56^{bright}$ ($CD56^{hi} CD16$).

5.3. Estudio de la proliferación y secreción de citoquinas tras estimulación de linfocitos T los diferentes haplotipos

CD6 es un miembro de la familia *scavenger* y las dos isoformas del CD6 parecen unir bacterias y por tanto, pueden actuar en una fase de infección temprana supuestamente asociada con la aparición de EM, aunque el mecanismo preciso todavía se desconoce. Los estudios funcionales usando anticuerpos monoclonales dirigidos contra el CD6 han demostrado su papel en la activación y proliferación de LTs²⁴⁵⁻²⁴⁸ y en la regulación de la expresión de fosfoproteínas intracelulares y la producción de citoquinas proinflamatorias²⁴⁹. Estudios de unión con lipopolisacáridos y anticuerpos monoclonales anti-CD6 muestran fosforilación de ERK1/2 y p358 MAK, indicando la activación de la vía MAPK^{250, 251}. La expresión diferencial de CD6 dependiendo del haplotipo presente apunta a un posible efecto de este gen en la función de los LTs. Por lo tanto, realizamos estudios adicionales para identificar diferencias en la proliferación y en la producción de citoquinas. Se estudiaron LTs estimulados policlonalmente para determinar si la coestimulación del CD6 usando anticuerpos específicos tenía un efecto en la proliferación y en la producción de citoquinas, relacionado con el haplotipo (Figura 17). Para este propósito, las PBMCs se estimularon con anti-CD3 en presencia o ausencia de anticuerpo monoclonal anti-CD6. Los estudios de proliferación y en los niveles de secreción de IFN- γ e IL-17A de citoquinas muestran una tendencia hacia una proliferación reducida en el haplotipo de riesgo (CT) cuando lo comparamos con el haplotipo protector (CC). Estos descubrimientos están ligados con los descubrimientos de Kofler *et al.*²³⁹, que observaron una proliferación celular reducida entre

los donantes con el genotipo de riesgo rs17824933GG. Además, se observaron niveles más altos de citoquina IL-17A con el haplotipo protector (CC) que con el haplotipo de riesgo (CT), de forma similar a lo observado en el estudio de proliferación (aunque no significativos).

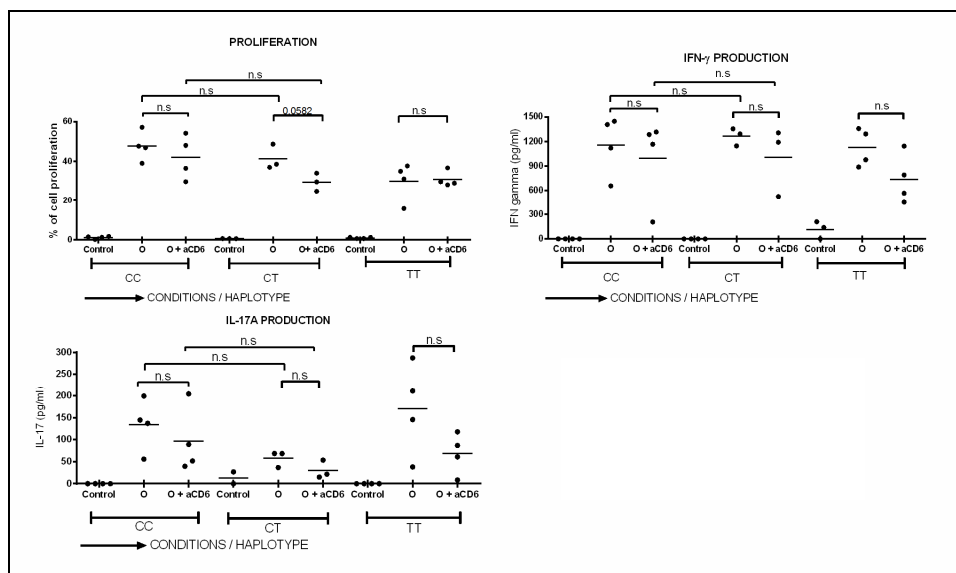


Figura 17. Estudio de la proliferación y secreción de citoquinas tras estimulación de LTs con los diferentes haplotipos rs11230563-rs2074225. Arriba izquierda) Proliferación de LTs, arriba derecha) Producción de IFN- γ ; y abajo izquierda) Producción de IL-17A. Control: células no estimuladas, O: células estimuladas con anticuerpo OKT3, aCD6: células coestimuladas con CD6, n.s.: no significativo.

Debido a que la adición de anti-CD6 al cultivo no modificó significativamente la tasa de proliferación y la producción de citoquinas, se podría inferir que el anticuerpo monoclonal anti-CD6 podría conducir a un bloqueo generalizado de la proliferación y de la producción de citoquinas y no es dependiente de SNPs. Las tendencias de IL-17A, aunque poco concluyentes, sugieren que la variación alélica del *CD6* se podría asociar con una secreción alterada de citoquinas de linfocitos Th17. Dada la capacidad del CD6 de unirse a bacterias/lipopolisacáridos²⁵⁰, y la evidencia de un papel de la IL-17A en mediar la protección contra varios patógenos^{252, 253}, esto podría ser relevante en el concepto de agentes

infecciosos actuando como desencadenantes de la EM. Alternativamente, la observación de niveles más altos de IL-17A entre aquellos con el haplotipo protector podría indicar un papel protector/anti-inflamatorio de la IL-17A ya que se ha demostrado que esta citoquina tiene la capacidad de proteger contra el desarrollo de uveítis autoinmune y colitis ulcerosa^{254, 255}. Sin embargo, esta afirmación contrasta con el rol descrito de los linfocitos Th17 en la EM^{30, 256, 257} y necesitaría una clarificación adicional.

6. Influencia de la delección en *LILRA3* en el riesgo a desarrollar EM

Nos propusimos estudiar la influencia de la delección de *LILRA3* en la susceptibilidad a EM en un número elevado de individuos de 3 localizaciones diferentes en España (Barcelona, Madrid y Andalucía) genotipadas para tal fin y posteriormente realizamos un meta-análisis con todos los datos disponibles hasta la fecha incluidos los previamente publicados. En total, nuestro estudio con 6.269 individuos (3.200 pacientes con EM y 3.069 controles pareados) no mostró una diferencia estadísticamente significativa cuando se comparó la distribución de la delección de *LILRA3* en pacientes con EM y controles [$P = 0,25$; OR (IC 95%) = 1,07 (0,95-1,19)], aunque en el contexto de una moderada evidencia de heterogeneidad ($I^2 = 32\%$) (Figura 18). Después de un análisis de sensibilidad se observó que la exclusión de la colección identificada como “2009-España (Madrid)” estudiada por Ordoñez *et al.*¹⁴⁰, consiguió reducir sustancialmente dicha heterogeneidad en el meta-análisis total, pero sin mostrar una asociación significativa ($P = 0,49$). La reducción máxima de la heterogeneidad se consiguió cuando el estudio con el impacto más alto en la susceptibilidad a EM se retiró del meta-análisis, apoyando la hipótesis de que el efecto observado anteriormente se debiera a una fuerte señal en esta colección. La asociación de la forma delecionada del gen *LILRA3* con la susceptibilidad a EM se había estudiado repetidamente¹³⁸⁻¹⁴¹. Sin embargo, los resultados no eran totalmente concordantes y era difícil extraer una conclusión firme en lo que respecta al impacto de este gen en EM. Cuatro estudios previamente publicados habían estudiado la asociación de la delección de *LILRA3* con la susceptibilidad a EM en 5 colecciones de datos individuales. Los estudios en poblaciones alemana¹³⁹ y española¹⁴⁰ mostraron una asociación significativa de la delección *LILRA3* con EM remitente (definida en su estudio por las formas EMRR y EMSP), siendo la intensidad de la asociación más fuerte en el estudio con población española. En este punto, hay que destacar que una colección de pacientes con EM de origen francés mostraba una alta frecuencia de la delección en *LILRA3* (46,1%), sin embargo, en dicho estudio no se incluían controles étnicamente pareados¹³⁹ y no pudo ser incluida en nuestro meta-análisis. Por el contrario, dos estudios llevados a cabo en población finlandesa¹³⁸ y polaca¹⁴¹ no pudieron replicar esta asociación. Por lo tanto, un sesgo de publicación no parece estar presente en el estudio de este gen. El meta-análisis combinando las colecciones mencionadas mostró un asociación estadísticamente significativa de la delección de *LILRA3* con la susceptibilidad a

EM [$P = 0,03$; OR (IC 95%) = 1,19 (1,02-1,40)], con evidencia modesta de heterogeneidad ($I^2 = 25\%$). Así, consideramos que el pequeño tamaño de las cohortes podría explicar la inconsistencia en los resultados observados en las diferentes publicaciones, pero teniendo en cuenta nuestros resultados parece que la delección de *LILRA3* por sí misma no parece alterar la susceptibilidad a EM.

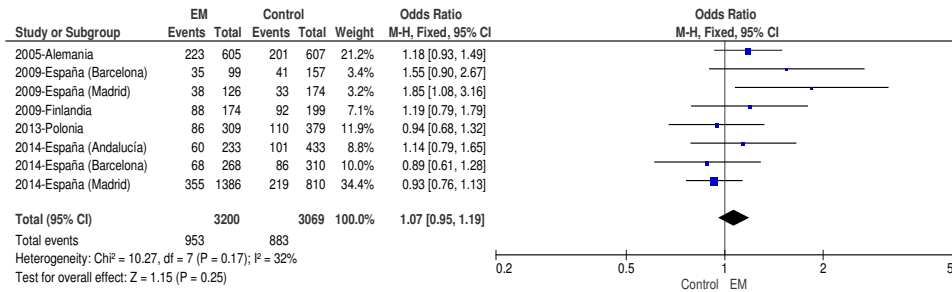


Figura 18. Meta-análisis de efectos aleatorios estudiando la asociación de la delección *LILRA3* con el riesgo a EM en poblaciones de ascendencia europea. ORs específicas de estudio (cuadrados) e IC al 95% (líneas) calculadas para cada colección individual. La OR conjunta (rombo) e IC al 95% se calculó combinado todas las bases de datos.

Posteriormente, con aquellas colecciones que incluían datos de los genotipos de *LILRA3* y *HLA-DRB1*15:01*, realizamos un meta-análisis adicional para estudiar si existía un efecto sinérgico entre *HLA-DRB1*15:01* y la delección *LILRA3* para conferir riesgo a EM como se describió anteriormente por Ordonez *et al.*¹⁴⁰, pero no se observó una interacción epistática significativa [$P = 0,36$, OR (IC 95%) = 1,09 (0,91-1,30)] (Figura 19). Para estudiar si el efecto de la delección de *LILRA3* podría restringirse a los sujetos que carecían del alelo *HLA-DRB1*15:01*, el anterior análisis se llevó a cabo en este grupo de pacientes, pero tampoco se observó una diferencia significativa [$P = 0,74$, OR (IC 95%) = 1,03 (0,88-1,20)].

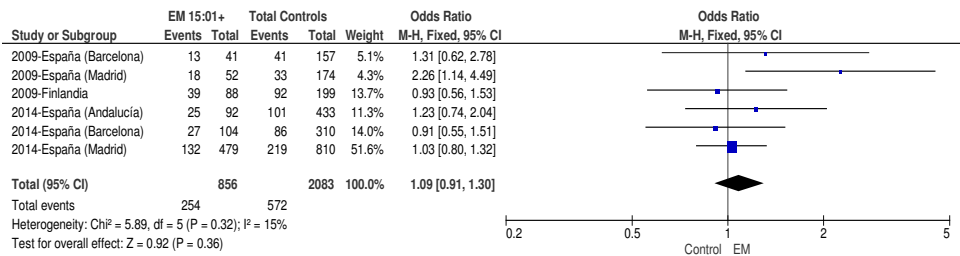


Figura 19. Meta-análisis de efectos aleatorios incluyendo los estudios de asociación de la delección de *LILRA3* con el riesgo a EM en portadores del alelo *HLA-DRB1*15:01* en poblaciones descendientes de europeos. ORs específicas de estudio (cuadrados) e IC al 95% (líneas) calculadas para cada colección individual. La OR conjunta (rombo) e IC al 95% se combinaron calculando todas las bases de datos.

Además, se estudió la interacción epistática después de la estratificación por género para evaluar si la delección de *LILRA3* presentaba un posible sesgo en el riesgo a conferir EM, como se había descrito por Du *et al.* en la susceptibilidad a AR¹⁷³. Así, se realizó un análisis adicional estratificando los sujetos doble positivos para la delección de *LILRA3* y *HLA-DRB1*15:01* en función del sexo, pero no se observó asociación [$P = 0,39$; OR (IC 95%) = 1,09 (0,90-1,31)] con una homogeneidad total entre las colecciones ($I^2 = 0\%$) (Figura 20). Al considerar los grupos de hombres y mujeres de forma individual se obtuvieron resultados similares [hombres: $P = 0,76$, OR (IC 95%) = 1,05 (0,76-1,47) y mujeres: $P = 0,40$, OR (IC 95%) = 1,10 (0,88-1,38)], aunque hay que considerar que se excluyeron dos de las colecciones iniciales, ya que carecían de la información detallada sobre el género y la presencia del alelo *HLA-DRB1*15:01*. Tampoco se observó una diferencia en función del sexo cuando se realizó un meta-análisis de los portadores de la delección de *LILRA3* que carecían del alelo *HLA-DRB1*15:01* (datos no mostrados). Por lo tanto, las estratificaciones considerando el alelo *HLA-DRB1*15:01* por sí mismo o incluyendo también el género de los individuos no mostró una asociación global de la delección de *LILRA3* con el riesgo a EM, demostrando que una estratificación más rigurosa de los pacientes no explica los resultados previos.

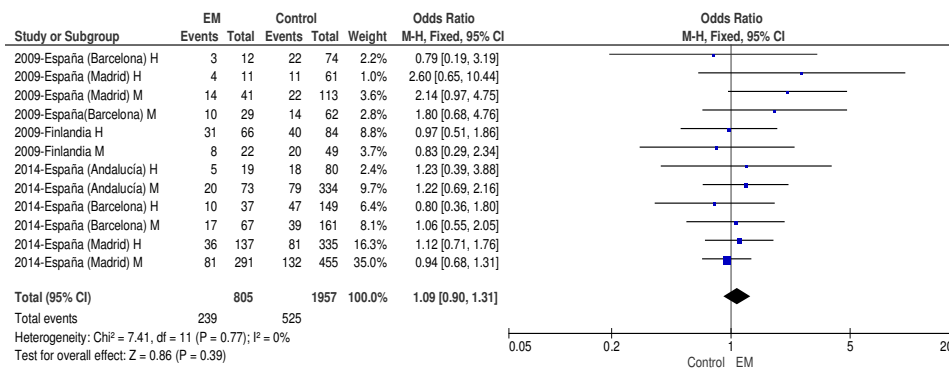


Figura 20. Meta-análisis de efectos aleatorios de los estudios de asociación de la delección *LILRA3* con el riesgo a EM en portadores del *HLA-DRB1*15:01* de poblaciones de ascendencia europea estratificados por sexo. ORs específicas de estudio (cuadrados) e IC al 95% (líneas) calculadas para cada colección individual. La OR conjunta (rombo) e IC al 95% se combinaron calculando todas las bases de datos. EM 15:01+: portadores de la delección de *LILRA3* y del alelo *HLA-DRB1*15:01*, H: hombre, M: mujer.

El papel de la genética en la susceptibilidad a EM es altamente aceptada y ha sido ampliamente descrita en la literatura. En la actualidad, los mayores avances en el estudio del componente genético se han conseguido a través de los GWAS, que buscan señales que expliquen la patogénesis de la EM analizando miles de pacientes y controles. Sin embargo, esta clase de estudios no exploran en variantes genéticas distintas a los SNPs y la heredabilidad que subyace a esta y otras enfermedades complejas persiste aún parcialmente inexplicada. En este sentido, como se ha comentado anteriormente, diversos estudios han tratado de evaluar la influencia de *LILRA3* en el riesgo a EM. Por otro lado, algunos estudios parecen apoyar el riesgo a otras enfermedades autoinmunes debido al gen *LILRA3* y este gen podría ser otro ejemplo del sustrato genético común entre este tipo de enfermedades^{173, 258, 259}. Sin embargo, estos estudios requieren validación en colecciones de individuos o poblaciones independientes debido a que los estudios que incluían un mayor número de individuos se realizaron en poblaciones asiáticas. La frecuencia de la delección de 6,7 kb en el gen *LILRA3* difiere entre los grupos étnicos con valores extremadamente altos en población asiática (0,56-0,84) y valores mucho más bajos en europeos (0,17) o africanos (0,10)²⁶⁰. Es

bien conocido que las poblaciones asiáticas y caucásicas muestran variantes de riesgo específicas debido a la diferente deriva genética y los diferentes patrones de LD resultantes.

Nuestros resultados son concordantes con los encontrados en psoriasis, que mostraban una ausencia de asociación entre la delección en el gen *LILRA3* y la enfermedad en población europea ²⁶¹, aunque este descubrimiento también requiere una replicación adicional. Toda esta información en su conjunto parece señalar que *LILRA3* no subyace a la patogénesis de la EM.

Estado actual y posibles aproximaciones en el estudio de los factores genéticos relacionados con la EM

Los mecanismos patofisiológicos de la EM probablemente son dinámicos en el tiempo y es difícil determinar si los efectos favorables del tratamiento serán sostenidos a largo plazo debido al curso crónico e impredecible de la enfermedad ¹. La EM se considera una enfermedad multifactorial, en la que tanto factores ambientales como genéticos actúan promoviendo la aparición de la enfermedad. Este trabajo ha intentado descifrar parte del componente genético que permanece aún inexplorado en EM. Nuestros esfuerzos proporcionan evidencia significativa de asociación de 9 loci con EM. Entre estos se incluyen: *TMEM39A* [rs1132200, $P = 0,01$; OR (IC 95%) = 0,84 (0,75-0,93)], *IL12B* [rs6887695; $P = 0,03$; OR (IC 95%) = 1,09 (1,01-1,17)], *CBLB* [rs9657904; $P = 0,01$; OR (IC 95%) = 0,89 (0,81-0,97)], *MANBA* [rs228614; $P = 2,4 \times 10^{-6}$; OR (IC 95%) = 0,91 (0,87-0,95)], *CXCR5* [rs630923; $P = 1,2 \times 10^{-4}$; OR (IC 95%) = 0,89 (0,84-0,95)], *SOX8* [rs2744148; $P = 1,8 \times 10^{-6}$; OR (IC 95%) = 1,14 (1,08-1,20)], *RPS6KBI* [rs180515; $P = 5,2 \times 10^{-7}$; OR (IC 95%) = 1,12 (1,07-1,17)], *ZBTB46* [rs6062314; $P = 4,3 \times 10^{-3}$; OR (IC 95%) = 0,90 (0,83-0,97)] y *SOCS1* [rs243324; $P = 6 \times 10^{-5}$; OR (IC 95%) = 1,24 (1,05-1,46)]. Los loci asociados evidencian un efecto modesto, pero abren vías de conocimiento de la patogénesis de esta condición discapacitante.

Por otro lado, nuestro trabajo parece poner fin al debate que existía en la literatura respecto a la asociación de los SNPs rs7412 y rs429358 del gen *APOE* con EM, ya que con más de 29.000 sujetos estudiados no pudimos observar asociación estadísticamente significativa en ningún caso [$P = 0,057$; OR (IC 95%) = 1,07 (1,00-1,15) y $P = 0,29$; OR (IC 95%) = 0,95 (0,88-1,04), respectivamente]. Lo mismo ocurre con la asociación de la delección de *LILRA3* y la susceptibilidad a EM, también estudiada en este trabajo y con resultados previos inconsistentes descritos por otros autores ¹³⁸⁻¹⁴¹, donde no observamos asociación ni cuando consideramos solo la delección, ni cuando realizamos análisis adicionales considerando el alelo *HLA-DRB1*15:01* con o sin una estratificación adicional en función del género de los pacientes.

Estos resultados ponen de manifiesto varios aspectos. En primer lugar, enfatizan la necesidad de replicar señales sugerentes de asociación detectadas a través de los GWAS, ya que asociaciones espurias pueden alcanzar valores de P respetables y solo pueden

distinguirse después de un ejercicio de replicación. Así, los estudios de replicación se concentran en los pocos resultados que suponen asociación al borde de la significación estadística, para identificar verdaderas asociaciones que podrían haber pasado por alto en los primeros estudios ²⁰⁶, tal como ha sido el caso de los loci: *TMEM39A*, *CBLB*, *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* y *ZBTB46*, descritos en este trabajo.

En segundo lugar, nuestros resultados proporcionan evidencia adicional del papel del sistema inmunitario en la etiología de la EM. Por ejemplo, *SOCS1* identificado como locus de susceptibilidad a EM en este trabajo, al igual que *CD6*, *IRF5* e *IRF8* anteriormente descritos en otros estudios ^{116, 262, 263} entre otros, son parte de una creciente serie de genes de susceptibilidad a EM cuyos productos tienen papeles establecidos en la defensa frente a agentes patógenos infecciosos. Son necesarios esfuerzos adicionales para comprender cómo la compleja inmunopatogénesis de la EM da lugar a su heterogeneidad característica y obtener una mejor visión de los mecanismos moleculares que subyacen a la EM.

En tercer lugar, nuestros resultados apoyan la existencia de variantes genéticas que predisponen a múltiples enfermedades autoinmunes a pesar de sus manifestaciones clínicas claramente diferentes. Esto sugiere un sustrato genético común, que determina mecanismos moleculares comunes ^{124, 125}. Así, el gen *IL12B* se encontró asociado con EM, tal como se había observado en otras enfermedades de base inmunológica como la psoriasis ^{101, 102, 264}, la colitis ulcerosa ^{210, 211} y el LES ²¹². Por lo tanto, debido al gran solapamiento genético entre distintas enfermedades autoinmunes, un mejor conocimiento de la EM no solo contribuirá a un mejor conocimiento de la misma, sino que en parte también podrá contribuir a un mejor conocimiento de enfermedades con las que comparte dicho sustrato genético. Desde esta perspectiva, algunos loci parecen ser factores específicos de una única enfermedad, lo que evidencia que existen distintos mecanismos determinantes de la especificidad de la enfermedad ¹²⁷.

La mayor parte de los SNPs asociados al riesgo a padecer EM en este trabajo no identifican el gen o genes dentro de una región que confieren susceptibilidad a la enfermedad, porque son marcadores de haplotipos en los que reside la variante funcional y la variación genética etiológica permanece sin conocerse. Además, el lugar en la fisiología o en la evolución de la enfermedad sigue estando sin dilucidar para muchos de estos genes. Las aproximaciones para pasar de marcadores de susceptibilidad a genes causantes de enfermedad incluyen una combinación de secuenciación de las regiones que contienen las

señales de asociación y análisis posteriores que impliquen una priorización de los supuestos SNPs funcionales seleccionados mediante la integración de métodos epidemiológicos y bioinformáticos, un genotipado de replicación pormenorizado y la posterior verificación funcional tanto *in vivo* como *in vitro* de los genes diana. En este sentido, en este trabajo se realizó un *tagging* del locus del *CD6*, se estudiaron los SNPs marcadores de la región en una colección de individuos, se replicaron las señales más significativas en una colección con un mayor número de individuos y se estudió el efecto de estos haplotipos en la expresión de *CD6* en superficie y la secreción de citoquinas. El análisis mostró patrones de expresión significativamente diferentes de *CD6* en los distintos subtipos celulares: células $CD4^+$ vírgenes, $P = 0,0001$; células $CD8^+$ vírgenes, $P < 0,0001$; células de memoria central $CD4^+$ y $CD8^+$, $P = 0,01$ y $0,05$, respectivamente; y células NKT, $P = 0,02$; con el haplotipo protector mostrando una alta expresión de *CD6*. Sin embargo, no se observaron cambios significativos ni en células NK, ni tampoco en células T de memoria efectoras y de memoria efectoras diferenciadas terminalmente. Nuestros resultados revelan que este nuevo haplotipo de riesgo del *CD6* asociado a EM modifica significativamente la expresión de *CD6* en células $CD4^+$ y $CD8^+$ ¹⁶⁴.

En los próximos años, la identificación de las consecuencias funcionales precisas de estas variantes de riesgo, cuyo propósito final es comprender cómo una combinación de múltiples variantes de riesgo genético puede perturbar sistemas moleculares en tipos celulares definidos, plantearán un considerable reto en la investigación de la EM. En este sentido, otros estudios han abierto nuevas vías de investigación en este campo. Por ejemplo, Fairfax *et al.* ²⁶⁵ y Lee *et al.* ²⁶⁶ estudiaron la expresión génica a nivel de genoma completo en monocitos y CDs que fueron expuestos a estímulos inflamatorios comparadas con células no estimuladas. Se observó que mientras un grupo de variantes genéticas afecta los niveles de expresión génica en células no estimuladas, otras variantes solo afectan los niveles de expresión génica en aquellas células que estimuladas por una señal inflamatoria. Así se identificaron muchas variantes que no se habían asociado previamente con cambios en la expresión génica.

La aproximación llevada a cabo en estos dos estudios con distintos tipos de células inmunes y también con muestras de pacientes con enfermedades inmunorrelacionadas, contribuirá a mejorar nuestra comprensión de cómo la variación genética humana afecta a la respuesta inmune y proporcionará hallazgos en los mecanismos subyacentes a la asociación

de variantes genéticas particulares con la enfermedad ²⁶⁷. La gran resolución del mapeo genético, junto con bases de datos genómicos a gran escala tales como ENCODE han revolucionado nuestra capacidad para formular hipótesis fácilmente testables ²⁶⁸, y en el futuro se espera que se definan papeles funcionales para muchas de estas variantes que predisponen a la enfermedad.

Aunque los loci identificados promueven aparentemente efectos modestos en la predisposición a EM, en la mayor parte de los casos no se ha considerado la potencial existencia de efectos epistáticos y epigenéticos, ni las posibles interacciones existentes entre las señales identificadas y factores ambientales. Tampoco se ha tenido en cuenta sistemáticamente hasta ahora el efecto que pudieran tener las variantes estudiadas en los distintos subgrupos de pacientes, lo que podría conducir a una mejor definición de su impacto real para comprender la causa y la progresión de la EM, como ya se ha probado para ciertos genes de susceptibilidad a EM ²⁶⁹.

La completa interpretación de estos datos genéticos requiere una comprensión clara de cómo estos nuevos genes de riesgo están regulados, algo que de acuerdo a los estudios realizados resultará complejo. Maurano *et al.* ²⁷⁰ observó que el 41% de los SNPs asociados con sitios de hipersensibilidad a DNasa regulan genes a más de 250 kb, y estos genes no son siempre los más cercanos, mostrando que los genes diana pueden estar más allá de las fronteras del LD incluso en cromosomas diferentes. Así, la localización de los genes diana con respecto al SNP marcador supone un reto.

El descubrimiento de las bases genéticas de la EM tiene un claro propósito para la biomedicina, porque nos permitirá seleccionar biomarcadores de evolución y de respuesta a tratamientos. Se presupone que varios procesos contribuyen secuencialmente o simultáneamente a la patofisiología de la EM y unos biomarcadores indicativos de estos procesos serían de gran utilidad. Nos permitirían realizar el diagnóstico y la identificación de los estados y subcategorías de la EM, ya que es importante discriminar aquellos pacientes en los que el curso clínico va a ser más severo, para tomar a la mayor brevedad posible las acciones terapéuticas que se consideren oportunas. Actualmente el curso de la EM en un paciente individual es en gran medida impredecible ¹ y la precisión de los modelos de riesgo genético será el factor determinante para su potencial uso en la práctica clínica en el futuro ²⁷¹. El valor individual de los SNP asociados a EM para desarrollar modelos predictivos actualmente es limitado, pero se espera desarrollar en el futuro modelos de múltiples

variantes genéticas. La predicción del desarrollo de la enfermedad basada en un panel de alelos de riesgo será particularmente importante en familias con una historia de EM o de otras enfermedades autoinmunes.

Descubrir todos los factores de riesgo genético de esta enfermedad compleja es crucial para proporcionar claves para establecer de forma más precisa los mecanismos moleculares subyacentes. Supuestamente, los productos codificados por los genes que predisponen a una enfermedad se relacionan mediante redes funcionales interconectadas, por lo que dilucidar las bases genéticas identificará en última estancia rutas biológicas importantes que podrían representar dianas adecuadas para el desarrollo de nuevos fármacos. La identificación de rutas asociadas con alelos de riesgo afectará las decisiones en el tratamiento de los pacientes con EM y además ofrece importantes oportunidades para alcanzar una medicina personalizada ²⁷². Asimismo, posibilitará la evaluación de nuevas terapias, ya que la identificación de los polimorfismos de riesgo ayudará a dilucidar las alteraciones en la señalización involucradas en el desarrollo de la EM y puede ayudar a seleccionar nuevas aproximaciones terapéuticas.

Los estudios incluidos en este trabajo se diseñaron con el propósito de explicar, al menos en parte, la heredabilidad perdida en EM. En los últimos años, se ha avanzado enormemente en el conocimiento de la base genética de la EM debido a la introducción de nuevas técnicas tanto metodológicas como analíticas y a la conciencia de que los grandes avances requieren extensas colaboraciones nacionales e internacionales y consorcios tales como el IMSGC. Este trabajo ejemplifica muy bien este último aspecto, ya que las conclusiones que se emiten son posibles gracias a la cooperación entre los distintos centros dedicados al estudio del componente genético asociado a EM tanto a nivel nacional como internacional.

Con todo aún estamos lejos de comprender la arquitectura genética completa de esta enfermedad compleja. Existe aún un largo camino por recorrer debido a que los actuales conocimientos y no son todavía clínicamente relevantes y en el futuro, el éxito dependerá del progreso de los estudios interdisciplinarios que crean mapas de genotipo-fenotipo para comprender mejor la patogénesis y las potenciales aproximaciones terapéuticas de la EM

Conclusiones

Los avances llevados a cabo mediante la realización del presente trabajo permiten extraer las siguientes conclusiones:

1. El SNP rs243324 del gen *SOCS1* es un marcador de riesgo a EM y esta asociación es independiente del gen de riesgo *CLEC16A* previamente descrito.
2. Los SNPs rs7412 y rs429358, localizados en el gen *APOE*, no parecen modificar la susceptibilidad a EM en poblaciones de ascendencia europea, como muestra nuestro estudio incluyendo 29.744 y 28.460 sujetos, respectivamente y combinando los resultados de genotipos *de novo* de múltiples poblaciones, datos imputados de GWAS y evidencia sistemáticamente cotejada de la literatura.
3. Los genes *IL12B*, gen candidato asociado con diversas enfermedades inmuno-mediadas; y *CBLB*, gen asociado inicialmente en la población italiana, se asocian al riesgo a presentar EM.
4. Los loci que engloban el gen *TMEM39A*, identificado previamente en un estudio de replicación del primer GWAS en EM y los genes *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* y *ZBTB46* previamente descritos al borde de la significación tras análisis por GWAS, se confirman en este estudio como loci de susceptibilidad a EM.
5. El haplotipo CT formado por los SNPs rs11230563 y rs2074225, del gen *CD6*, confiere riesgo a EM y modifica significativamente la expresión de *CD6* en células *CD4⁺* y *CD8⁺*.
6. La delección en el gen *LILRA3* no está asociada a EM por sí misma ni cuando se tienen en cuenta adicionalmente factores como el alelo *HLA*DRB1*15:01* o el género.

Bibliografía

1. Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG. Multiple sclerosis. *N Engl J Med* 2000; **343**: 938-952.
2. MSIF. Atlas of MS 2013: Mapping multiple sclerosis around the world. 2013.
3. Confavreux C, Vukusic S. Natural history of multiple sclerosis: a unifying concept. *Brain* 2006; **129**: 606-616.
4. Pugliatti M, Rosati G, Carton H, Riise T, Drulovic J, Vecsei L *et al.* The epidemiology of multiple sclerosis in Europe. *Eur J Neurol* 2006; **13**: 700-722.
5. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology* 1996; **46**: 907-911.
6. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ *et al.* Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology* 2014; **83**: 278-286.
7. Ebers GC, Daumer M. Natural history of MS. *Eur J Neurol* 2008; **15**: 881-882.
8. Kamm CP, Uitdehaag BM, Polman CH. Multiple Sclerosis: Current Knowledge and Future Outlook. *Eur Neurol* 2014; **72**: 132-141.
9. Compston A, Coles A. Multiple sclerosis. *Lancet* 2008; **372**: 1502-1517.
10. Alvarez-Cermeno JC, Villar LM. Multiple sclerosis: Oligoclonal bands--a useful tool to avoid MS misdiagnosis. *Nat Rev Neurol* 2013; **9**: 303-304.
11. Cortese I, McFarland HF. Chapter 65: Multiple Sclerosis. *Clinical Immunology: Principles and Practise* 2008: 963-976.
12. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC *et al.* New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983; **13**: 227-231.
13. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001; **50**: 121-127.
14. Zozulya AL, Wiendl H. The role of regulatory T cells in multiple sclerosis. *Nat Clin Pract Neurol* 2008; **4**: 384-398.
15. Nylander A, Hafler DA. Multiple sclerosis. *J Clin Invest* 2012; **122**: 1180-1188.
16. Sospedra M, Martin R. Immunology of multiple sclerosis. *Annu Rev Immunol* 2005; **23**: 683-747.

17. Steinman L, Zamvil SS. How to successfully apply animal studies in experimental allergic encephalomyelitis to research on multiple sclerosis. *Ann Neurol* 2006; **60**: 12-21.
18. Wucherpfennig KW, Strominger JL. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell* 1995; **80**: 695-705.
19. Munz C, Lunemann JD, Getts MT, Miller SD. Antiviral immune responses: triggers of or triggered by autoimmunity? *Nat Rev Immunol* 2009; **9**: 246-258.
20. Furtado GC, Marcondes MC, Latkowski JA, Tsai J, Wensky A, Lafaille JJ. Swift entry of myelin-specific T lymphocytes into the central nervous system in spontaneous autoimmune encephalomyelitis. *J Immunol* 2008; **181**: 4648-4655.
21. Zhang H, Podojil JR, Luo X, Miller SD. Intrinsic and induced regulation of the age-associated onset of spontaneous experimental autoimmune encephalomyelitis. *J Immunol* 2008; **181**: 4638-4647.
22. McQualter JL, Bernard CC. Multiple sclerosis: a battle between destruction and repair. *J Neurochem* 2007; **100**: 295-306.
23. Mikulkova Z, Praksova P, Stourac P, Bednarik J, Michalek J. Imbalance in T-cell and cytokine profiles in patients with relapsing-remitting multiple sclerosis. *J Neurol Sci* 2011; **300**: 135-141.
24. Duszczyszyn DA, Beck JD, Antel J, Bar-Or A, Lapierre Y, Gadag V *et al*. Altered naive CD4 and CD8 T cell homeostasis in patients with relapsing-remitting multiple sclerosis: thymic versus peripheral (non-thymic) mechanisms. *Clin Exp Immunol* 2006; **143**: 305-313.
25. Haak S, Croxford AL, Kreymborg K, Heppner FL, Pouly S, Becher B *et al*. IL-17A and IL-17F do not contribute vitally to autoimmune neuro-inflammation in mice. *J Clin Invest* 2009; **119**: 61-69.
26. Gocke AR, Cravens PD, Ben LH, Hussain RZ, Northrop SC, Racke MK *et al*. T-bet regulates the fate of Th1 and Th17 lymphocytes in autoimmunity. *J Immunol* 2007; **178**: 1341-1348.
27. Stromnes IM, Cerretti LM, Liggitt D, Harris RA, Goverman JM. Differential regulation of central nervous system autoimmunity by T(H)1 and T(H)17 cells. *Nat Med* 2008; **14**: 337-342.
28. Kebir H, Ifergan I, Alvarez JI, Bernard M, Poirier J, Arbour N *et al*. Preferential recruitment of interferon-gamma-expressing TH17 cells in multiple sclerosis. *Ann Neurol* 2009; **66**: 390-402.

29. Tzartos JS, Friese MA, Craner MJ, Palace J, Newcombe J, Esiri MM *et al.* Interleukin-17 production in central nervous system-infiltrating T cells and glial cells is associated with active disease in multiple sclerosis. *Am J Pathol* 2008; **172**: 146-155.
30. Brucklacher-Waldert V, Stuermer K, Kolster M, Wolthausen J, Tolosa E. Phenotypical and functional characterization of T helper 17 cells in multiple sclerosis. *Brain* 2009; **132**: 3329-3341.
31. Liao JJ, Huang MC, Goetzl EJ. Cutting edge: Alternative signaling of Th17 cell development by sphingosine 1-phosphate. *J Immunol* 2007; **178**: 5425-5428.
32. Buc M. Role of regulatory T cells in pathogenesis and biological therapy of multiple sclerosis. *Mediators Inflamm* 2013; **2013**: 963748.
33. Babbe H, Roers A, Waisman A, Lassmann H, Goebels N, Hohlfeld R *et al.* Clonal expansions of CD8(+) T cells dominate the T cell infiltrate in active multiple sclerosis lesions as shown by micromanipulation and single cell polymerase chain reaction. *J Exp Med* 2000; **192**: 393-404.
34. Bitsch A, Schuchardt J, Bunkowski S, Kuhlmann T, Bruck W. Acute axonal injury in multiple sclerosis. Correlation with demyelination and inflammation. *Brain* 2000; **123** (Pt 6): 1174-1183.
35. Skulina C, Schmidt S, Dornmair K, Babbe H, Roers A, Rajewsky K *et al.* Multiple sclerosis: brain-infiltrating CD8+ T cells persist as clonal expansions in the cerebrospinal fluid and blood. *Proc Natl Acad Sci U S A* 2004; **101**: 2428-2433.
36. Johnson AJ, Suidan GL, McDole J, Pirko I. The CD8 T cell in multiple sclerosis: suppressor cell or mediator of neuropathology? *Int Rev Neurobiol* 2007; **79**: 73-97.
37. Huseby ES, Liggitt D, Brabb T, Schnabel B, Ohlen C, Goverman J. A pathogenic role for myelin-specific CD8(+) T cells in a model for multiple sclerosis. *J Exp Med* 2001; **194**: 669-676.
38. Medana I, Martinic MA, Wekerle H, Neumann H. Transection of major histocompatibility complex class I-induced neurites by cytotoxic T lymphocytes. *Am J Pathol* 2001; **159**: 809-815.
39. O'Connor KC, Appel H, Bregoli L, Call ME, Catz I, Chan JA *et al.* Antibodies from inflamed central nervous system tissue recognize myelin oligodendrocyte glycoprotein. *J Immunol* 2005; **175**: 1974-1982.
40. Qin Y, Duquette P, Zhang Y, Talbot P, Poole R, Antel J. Clonal expansion and somatic hypermutation of V(H) genes of B cells from cerebrospinal fluid in multiple sclerosis. *J Clin Invest* 1998; **102**: 1045-1050.

41. Obermeier B, Mentele R, Malotka J, Kellermann J, Kumpfel T, Wekerle H *et al.* Matching of oligoclonal immunoglobulin transcriptomes and proteomes of cerebrospinal fluid in multiple sclerosis. *Nat Med* 2008; **14**: 688-693.
42. Magliozzi R, Howell O, Vora A, Serafini B, Nicholas R, Puopolo M *et al.* Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. *Brain* 2007; **130**: 1089-1104.
43. Hauser SL, Waubant E, Arnold DL, Vollmer T, Antel J, Fox RJ *et al.* B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. *N Engl J Med* 2008; **358**: 676-688.
44. Weiner HL. The challenge of multiple sclerosis: how do we cure a chronic heterogeneous disease? *Ann Neurol* 2009; **65**: 239-248.
45. Gandhi R, Laroni A, Weiner HL. Role of the innate immune system in the pathogenesis of multiple sclerosis. *J Neuroimmunol* 2010; **221**: 7-14.
46. Rio J, Comabella M, Montalban X. Multiple sclerosis: current treatment algorithms. *Curr Opin Neurol* 2011; **24**: 230-237.
47. Maurer M, Dachsel R, Domke S, Ries S, Reifschneider G, Friedrich A *et al.* Health care situation of patients with relapsing-remitting multiple sclerosis receiving immunomodulatory therapy: a retrospective survey of more than 9000 German patients with MS. *Eur J Neurol* 2011; **18**: 1036-1045.
48. Rio J, Tintore M, Sastre-Garriga J, Nos C, Castillo J, Tur C *et al.* Change in the clinical activity of multiple sclerosis after treatment switch for suboptimal response. *Eur J Neurol* 2012; **19**: 899-904.
49. Luessi F, Siffrin V, Zipp F. Neurodegeneration in multiple sclerosis: novel treatment strategies. *Expert Rev Neurother* 2012; **12**: 1061-1076; quiz 1077.
50. Aharoni R. The mechanism of action of glatiramer acetate in multiple sclerosis and beyond. *Autoimmun Rev* 2013; **12**: 543-553.
51. Steinman L. Blocking adhesion molecules as therapy for multiple sclerosis: natalizumab. *Nat Rev Drug Discov* 2005; **4**: 510-518.
52. Ingwersen J, Aktas O, Kuery P, Kieseier B, Boyko A, Hartung HP. Fingolimod in multiple sclerosis: mechanisms of action and clinical efficacy. *Clin Immunol* 2012; **142**: 15-24.
53. Cruz VT, Fonseca J. Central effects of fingolimod. *Rev Neurol* 2014; **59**: 121-128.
54. Minagar A, Alexander JS, Sahraian MA, Zivadinov R. Alemtuzumab and multiple sclerosis: therapeutic application. *Expert Opin Biol Ther* 2010; **10**: 421-429.

55. Bielekova B. Daclizumab therapy for multiple sclerosis. *Neurotherapeutics* 2013; **10**: 55-67.
56. Barun B, Bar-Or A. Treatment of multiple sclerosis with anti-CD20 antibodies. *Clin Immunol* 2012; **142**: 31-37.
57. Ramagopalan SV, Dobson R, Meier UC, Giovannoni G. Multiple sclerosis: risk factors, prodromes, and potential causal pathways. *Lancet Neurol* 2010; **9**: 727-739.
58. Alonso A, Hernan MA. Temporal trends in the incidence of multiple sclerosis: a systematic review. *Neurology* 2008; **71**: 129-135.
59. Koch-Henriksen N, Sorensen PS. The changing demographic pattern of multiple sclerosis epidemiology. *Lancet Neurol* 2010; **9**: 520-532.
60. Ascherio A, Munger KL, Simon KC. Vitamin D and multiple sclerosis. *Lancet Neurol* 2010; **9**: 599-612.
61. Wingerchuk DM. Smoking: effects on multiple sclerosis susceptibility and disease progression. *Ther Adv Neurol Disord* 2012; **5**: 13-22.
62. Ascherio A, Munger KL. Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Ann Neurol* 2007; **61**: 288-299.
63. Berer K, Mues M, Koutrolos M, Rasbi ZA, Boziki M, Johner C *et al.* Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. *Nature* 2011; **479**: 538-541.
64. Elian M, Nightingale S, Dean G. Multiple sclerosis among United Kingdom-born children of immigrants from the Indian subcontinent, Africa and the West Indies. *J Neurol Neurosurg Psychiatry* 1990; **53**: 906-911.
65. Willer CJ, Dyment DA, Sadovnick AD, Rothwell PM, Murray TJ, Ebers GC. Timing of birth and risk of multiple sclerosis: population based study. *Bmj* 2005; **330**: 120.
66. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. *Am J Med Genet* 1988; **29**: 533-541.
67. Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A* 2003; **100**: 12877-12882.
68. Ebers GC, Bulman DE, Sadovnick AD, Paty DW, Warren S, Hader W *et al.* A population-based study of multiple sclerosis in twins. *N Engl J Med* 1986; **315**: 1638-1642.
69. Mumford CJ, Wood NW, Kellar-Wood H, Thorpe JW, Miller DH, Compston DA. The British Isles survey of multiple sclerosis in twins. *Neurology* 1994; **44**: 11-15.

70. Hansen T, Skytthe A, Stenager E, Petersen HC, Bronnum-Hansen H, Kyvik KO. Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. *Mult Scler* 2005; **11**: 504-510.
71. Sadovnick AD, Armstrong H, Rice GP, Bulman D, Hashimoto L, Paty DW *et al.* A population-based study of multiple sclerosis in twins: update. *Ann Neurol* 1993; **33**: 281-285.
72. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature* 1995; **377**: 150-151.
73. Ebers GC, Yee IM, Sadovnick AD, Duquette P. Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. *Ann Neurol* 2000; **48**: 927-931.
74. Sadovnick AD, Ebers GC, Dyment DA, Risch NJ. Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group. *Lancet* 1996; **347**: 1728-1730.
75. Ebers GC, Sadovnick AD, Dyment DA, Yee IM, Willer CJ, Risch N. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. *Lancet* 2004; **363**: 1773-1774.
76. Oksenberg JR, Baranzini SE, Sawcer S, Hauser SL. The genetics of multiple sclerosis: SNPs to pathways to pathogenesis. *Nat Rev Genet* 2008; **9**: 516-526.
77. Traherne JA. Human MHC architecture and evolution: implications for disease association studies. *Int J Immunogenet* 2008; **35**: 179-192.
78. Hauser SL, Oksenberg JR. The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. *Neuron* 2006; **52**: 61-76.
79. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. *Lancet* 1972; **1**: 1240-1241.
80. Terasaki PI, Park MS, Opelz G, Ting A. Multiple sclerosis and high incidence of a B lymphocyte antigen. *Science* 1976; **193**: 1245-1247.
81. Compston DA, Batchelor JR, McDonald WI. B-lymphocyte alloantigens associated with multiple sclerosis. *Lancet* 1976; **2**: 1261-1265.
82. Haines JL, Ter-Minassian M, Bazyk A, Gusella JF, Kim DJ, Terwedow H *et al.* A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. The Multiple Sclerosis Genetics Group. *Nat Genet* 1996; **13**: 469-471.
83. Stratmann T, Martin-Orozco N, Mallet-Designé V, Poirot L, McGavern D, Losyev G *et al.* Susceptible MHC alleles, not background genes, select an autoimmune T cell reactivity. *J Clin Invest* 2003; **112**: 902-914.

84. Wucherpfennig KW. The structural interactions between T cell receptors and MHC-peptide complexes place physical limits on self-nonself discrimination. *Curr Top Microbiol Immunol* 2005; **296**: 19-37.
85. Hoppenbrouwers IA, Hintzen RQ. Genetics of multiple sclerosis. *Biochim Biophys Acta* 2011; **1812**: 194-201.
86. Willer CJ, Dyment DA, Cherny S, Ramagopalan SV, Herrera BM, Morrison KM *et al.* A genome-wide scan in forty large pedigrees with multiple sclerosis. *J Hum Genet* 2007; **52**: 955-962.
87. Ebers GC, Kukay K, Bulman DE, Sadovnick AD, Rice G, Anderson C *et al.* A full genome search in multiple sclerosis. *Nat Genet* 1996; **13**: 472-476.
88. Sawcer S, Jones HB, Feakes R, Gray J, Smaldon N, Chataway J *et al.* A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet* 1996; **13**: 464-468.
89. Sawcer S, Ban M, Maranian M, Yeo TW, Compston A, Kirby A *et al.* A high-density screen for linkage in multiple sclerosis. *Am J Hum Genet* 2005; **77**: 454-467.
90. Sawcer S. The complex genetics of multiple sclerosis: pitfalls and prospects. *Brain* 2008; **131**: 3118-3131.
91. Becker KG, Simon RM, Bailey-Wilson JE, Freidlin B, Biddison WE, McFarland HF *et al.* Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci U S A* 1998; **95**: 9979-9984.
92. Becker KG. The common variants/multiple disease hypothesis of common complex genetic disorders. *Med Hypotheses* 2004; **62**: 309-317.
93. Schreiber S, Rosenstiel P, Albrecht M, Hampe J, Krawczak M. Genetics of Crohn disease, an archetypal inflammatory barrier disease. *Nat Rev Genet* 2005; **6**: 376-388.
94. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516-1517.
95. Wise LH, Lanchbury JS, Lewis CM. Meta-analysis of genome searches. *Ann Hum Genet* 1999; **63**: 263-272.
96. Vergheze PB, Castellano JM, Holtzman DM. Apolipoprotein E in Alzheimer's disease and other neurological disorders. *Lancet Neurol* 2011; **10**: 241-252.
97. Hoehe MR, Caenazzo L, Martinez MM, Hsieh WT, Modi WS, Gershon ES *et al.* Genetic and physical mapping of the human cannabinoid receptor gene to chromosome 6q14-q15. *New Biol* 1991; **3**: 880-885.

98. Ramil E, Sanchez AJ, Gonzalez-Perez P, Rodriguez-Antiguedad A, Gomez-Lozano N, Ortiz P *et al.* The cannabinoid receptor 1 gene (CNR1) and multiple sclerosis: an association study in two case-control groups from Spain. *Mult Scler* 2010; **16**: 139-146.
99. Pertwee RG. Cannabinoids and multiple sclerosis. *Pharmacol Ther* 2002; **95**: 165-174.
100. Tsunemi Y, Saeki H, Nakamura K, Sekiya T, Hirai K, Fujita H *et al.* Interleukin-12 p40 gene (IL12B) 3'-untranslated region polymorphism is associated with susceptibility to atopic dermatitis and psoriasis vulgaris. *J Dermatol Sci* 2002; **30**: 161-166.
101. Cargill M, Schrodi SJ, Chang M, Garcia VE, Brandon R, Callis KP *et al.* A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am J Hum Genet* 2007; **80**: 273-290.
102. Wu Y, Lu Z, Chen Y, Xue F, Chen X, Zheng J. Replication of association between interleukin-23 receptor (IL-23R) and its ligand (IL-12B) polymorphisms and psoriasis in the Chinese Han population. *Hum Immunol* 2010; **71**: 1255-1258.
103. Fisher SA, Tremelling M, Anderson CA, Gwilliam R, Bumpstead S, Prescott NJ *et al.* Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 2008; **40**: 710-712.
104. Danoy P, Pryce K, Hadler J, Bradbury LA, Farrar C, Pointon J *et al.* Association of variants at 1q32 and STAT3 with ankylosing spondylitis suggests genetic overlap with Crohn's disease. *PLoS Genet* 2010; **6**: e1001195.
105. International HapMap Consortium. The International HapMap Project. *Nature* 2003; **426**: 789-796.
106. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 2009; **106**: 9362-9367.
107. Panagiotou OA, Ioannidis JP. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int J Epidemiol* 2012; **41**: 273-286.
108. McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. *Hum Mol Genet* 2008; **17**: R156-165.

109. Sanna S, Pitzalis M, Zoledziewska M, Zara I, Sidore C, Murru R *et al.* Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat Genet* 2010; **42**: 495-497.
110. Corrado L, Bergamaschi L, Barizzzone N, Fasano ME, Guerini FR, Salvetti M *et al.* Association of the CBLB gene with multiple sclerosis: new evidence from a replication study in an Italian population. *J Med Genet* 2011; **48**: 210-211.
111. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL *et al.* Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007; **357**: 851-862.
112. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A *et al.* Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007; **39**: 1329-1337.
113. Comabella M, Craig DW, Camina-Tato M, Morcillo C, Lopez C, Navarro A *et al.* Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. *PLoS One* 2008; **3**: e3490.
114. Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, Broer L, Jafari N, Hillert J *et al.* Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. *Nat Genet* 2008; **40**: 1402-1403.
115. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009; **18**: 767-778.
116. De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT *et al.* Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 2009; **41**: 776-782.
117. Bahlo M, Booth D, Broadley S, (ANZgene). AaNZMSGC. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009; **41**: 824-828.
118. Nischwitz S, Cepok S, Kroner A, Wolf C, Knop M, Muller-Sarnowski F *et al.* Evidence for VAV2 and ZNF433 as susceptibility genes for multiple sclerosis. *J Neuroimmunol* 2010; **227**: 162-166.
119. Jakkula E, Leppa V, Sulonen AM, Varilo T, Kallio S, Kemppinen A *et al.* Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. *Am J Hum Genet* 2010; **86**: 285-291.

120. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011; **476**: 214-219.
121. Patsopoulos NA, Esposito F, Reischl J, Lehr S, Bauer D, Heubach J *et al.* Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* 2011; **70**: 897-912.
122. Matesanz F, Gonzalez-Perez A, Lucas M, Sanna S, Gayan J, Urcelay E *et al.* Genome-wide association study of multiple sclerosis confirms a novel locus at 5p13.1. *PLoS One* 2012; **7**: e36140.
123. Martinelli-Boneschi F, Esposito F, Brambilla P, Lindstrom E, Lavorgna G, Stankovich J *et al.* A genome-wide association study in progressive multiple sclerosis. *Mult Scler* 2012; **18**: 1384-1394.
124. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet* 2009; **10**: 43-55.
125. Baranzini SE. The genetics of autoimmune diseases: a networked perspective. *Curr Opin Immunol* 2009; **21**: 596-605.
126. Lettre G, Rioux JD. Autoimmune diseases: insights from genome-wide association studies. *Hum Mol Genet* 2008; **17**: R116-121.
127. Gregersen PK, Olsson LM. Recent advances in the genetics of autoimmune disease. *Annu Rev Immunol* 2009; **27**: 363-391.
128. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ *et al.* Finding the missing heritability of complex diseases. *Nature* 2009; **461**: 747-753.
129. Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kempainen A, Cotsapas C *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet* 2013; **45**: 1353-1360.
130. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565-569.
131. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet* 2012; **13**: 135-145.
132. Hunt KA, Mistry V, Bockett NA, Ahmad T, Ban M, Barker JN *et al.* Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. *Nature* 2013; **498**: 232-235.
133. Ramagopalan SV, Dyment DA, Cader MZ, Morrison KM, Disanto G, Morahan JM *et al.* Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Ann Neurol* 2011; **70**: 881-886.

134. Ban M, Goris A, Lorentzen AR, Baker A, Mihalova T, Ingram G *et al.* Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. *Eur J Hum Genet* 2009; **17**: 1309-1313.
135. Barizzzone N, Pauwels I, Luciano B, Franckaert D, Guerini FR, Cosemans L *et al.* No evidence for a role of rare CYP27B1 functional variations in multiple sclerosis. *Ann Neurol* 2013; **73**: 433-437.
136. Norman PJ, Carey BS, Stephens HA, Vaughan RW. DNA sequence variation and molecular genotyping of natural killer leukocyte immunoglobulin-like receptor, LILRA3. *Immunogenetics* 2003; **55**: 165-171.
137. Torkar M, Haude A, Milne S, Beck S, Trowsdale J, Wilson MJ. Arrangement of the ILT gene cluster: a common null allele of the ILT6 gene results from a 6.7-kbp deletion. *Eur J Immunol* 2000; **30**: 3655-3662.
138. Bonetti A, Koivisto K, Pirttila T, Elovaara I, Reunanen M, Laaksonen M *et al.* A follow-up study of chromosome 19q13 in multiple sclerosis susceptibility. *J Neuroimmunol* 2009; **208**: 119-124.
139. Koch S, Goedde R, Nigmatova V, Epplen JT, Muller N, de Seze J *et al.* Association of multiple sclerosis with ILT6 deficiency. *Genes Immun* 2005; **6**: 445-447.
140. Ordonez D, Sanchez AJ, Martinez-Rodriguez JE, Cisneros E, Ramil E, Romo N *et al.* Multiple sclerosis associates with LILRA3 deletion in Spanish patients. *Genes Immun* 2009; **10**: 579-585.
141. Wisniewski A, Wagner M, Nowak I, Bilinska M, Pokryszko-Dragan A, Jasek M *et al.* 6.7-kbp deletion in LILRA3 (ILT6) gene is associated with later onset of the multiple sclerosis in a Polish population. *Hum Immunol* 2013; **74**: 353-357.
142. Koch MW, Metz LM, Kovalchuk O. Epigenetics and miRNAs in the diagnosis and treatment of multiple sclerosis. *Trends Mol Med* 2013; **19**: 23-30.
143. Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 2010; **465**: 721-727.
144. Grossniklaus U, Kelly WG, Ferguson-Smith AC, Pembrey M, Lindquist S. Transgenerational epigenetic inheritance: how important is it? *Nat Rev Genet* 2013; **14**: 228-235.
145. Marian AJ. Elements of 'missing heritability'. *Curr Opin Cardiol* 2012; **27**: 197-201.
146. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A* 2012; **109**: 1193-1198.

147. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol* 2011; **29**: 493-525.
148. van der Mei IA, Ponsonby AL, Taylor BV, Stankovich J, Dickinson JL, Foote S *et al.* Human leukocyte antigen-DR15, low infant sibling exposure and multiple sclerosis: gene-environment interaction. *Ann Neurol* 2010; **67**: 261-265.
149. van der Mei IA, Simpson S, Jr., Stankovich J, Taylor BV. Individual and joint action of environmental factors and risk of MS. *Neurol Clin* 2011; **29**: 233-255.
150. Simon KC, van der Mei IA, Munger KL, Ponsonby A, Dickinson J, Dwyer T *et al.* Combined effects of smoking, anti-EBNA antibodies, and HLA-DRB1*1501 on multiple sclerosis risk. *Neurology* 2010; **74**: 1365-1371.
151. Ramagopalan SV, Maugeri NJ, Handunnetthi L, Lincoln MR, Orton SM, Dymment DA *et al.* Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D. *PLoS Genet* 2009; **5**: e1000369.
152. Hedstrom AK, Sundqvist E, Baarnhielm M, Nordin N, Hillert J, Kockum I *et al.* Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. *Brain* 2011; **134**: 653-664.
153. Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D *et al.* Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* 2009; **18**: 2078-2090.
154. Ramanujam R, Ramanujam S, Hillert J. Considerations for subgroups and phenocopies in complex disease genetics. *PLoS One* 2013; **8**: e71614.
155. Sawcer S, Franklin RJ, Ban M. Multiple sclerosis genetics. *Lancet Neurol* 2014.
156. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
157. Goris A, Walton A, Ban M, Dubois B, Compston A, Sawcer S. A Taqman assay for high-throughput genotyping of the multiple sclerosis-associated HLA-DRB1*1501 allele. *Tissue Antigens* 2008; **72**: 401-403.
158. Zivkovic M, Stankovic A, Dincic E, Popovic M, Popovic S, Raicevic R *et al.* The tag SNP for HLA-DRB1*1501, rs3135388, is significantly associated with multiple sclerosis susceptibility: cost-effective high-throughput detection by real-time PCR. *Clin Chim Acta* 2009; **406**: 27-30.

159. Lill CM, Liu T, Schjeide BM, Roehr JT, Akkad DA, Damotte V *et al.* Closing the case of APOE in multiple sclerosis: no association with disease risk in over 29 000 subjects. *J Med Genet* 2012; **49**: 558-562.
160. Varade J, Comabella M, Ortiz MA, Arroyo R, Fernandez O, Pinto-Medel MJ *et al.* Replication study of 10 genes showing evidence for association with multiple sclerosis: validation of TMEM39A, IL12B and CBLB [correction of CLBL] genes. *Mult Scler* 2012; **18**: 959-965.
161. Lill CM, Schjeide BM, Graetz C, Ban M, Alcina A, Ortiz MA *et al.* MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain* 2013; **136**: 1778-1782.
162. Vandenbroeck K, Alvarez J, Swaminathan B, Alloza I, Matesanz F, Urcelay E *et al.* A cytokine gene screen uncovers SOCS1 as genetic risk factor for multiple sclerosis. *Genes Immun* 2012; **13**: 21-28.
163. Ortiz MA, Núñez C, Ordóñez D, Alvarez-Cermeño JC, Martínez-Rodríguez JE, Sánchez AJ *et al.* Influence of the LILRA3 deletion on the risk of developing multiple sclerosis: a meta-analysis. *Enviado para su publicación* 2015.
164. Swaminathan B, Cuapio A, Alloza I, Matesanz F, Alcina A, Garcia-Barcina M *et al.* Fine mapping and functional analysis of the multiple sclerosis risk gene CD6. *PLoS One* 2013; **8**: e62376.
165. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559-575.
166. Koch S, Larbi A, Derhovanessian E, Ozcelik D, Naumova E, Pawelec G. Multiparameter flow cytometric analysis of CD4 and CD8 T cell subsets in young and old people. *Immun Ageing* 2008; **5**: 6.
167. Hamann D, Baars PA, Rep MH, Hooibrink B, Kerkhof-Garde SR, Klein MR *et al.* Phenotypic and functional separation of memory and effector human CD8+ T cells. *J Exp Med* 1997; **186**: 1407-1418.
168. Lanier LL, Le AM, Civin CI, Loken MR, Phillips JH. The relationship of CD16 (Leu-11) and Leu-19 (NKH-1) antigen expression on human peripheral blood NK cells and cytotoxic T lymphocytes. *J Immunol* 1986; **136**: 4480-4486.
169. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; **19**: 149-150.

170. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006; **38**: 209-213.
171. Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet* 2009; **10**: 318-329.
172. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263-265.
173. Du Y, Cui Y, Liu X, Hu F, Yang Y, Wu X *et al.* Contribution of functional LILRA3, but not nonfunctional LILRA3, to sex bias in susceptibility and severity of anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheumatol* 2014; **66**: 822-830.
174. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177-188.
175. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**: 2190-2191.
176. International Multiple Sclerosis Genetics Consortium (IMSGC). IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. *Genes Immun* 2010; **11**: 397-405.
177. Lundmark F, Duvefelt K, Iacobaeus E, Kockum I, Wallstrom E, Khademi M *et al.* Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* 2007; **39**: 1108-1113.
178. O'Doherty C, Kantarci O, Vandembroeck K. IL7RA polymorphisms and susceptibility to multiple sclerosis. *N Engl J Med* 2008; **358**: 753-754.
179. Vandembroeck K. *Cytokine Gene Polymorphisms in Multifactorial Conditions*: CRC Press: Boca Ratón, 2006.
180. Baranzini SE, Srinivasan R, Khankhanian P, Okuda DT, Nelson SJ, Matthews PM *et al.* Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis. *Brain* 2010; **133**: 2603-2611.
181. Byun E, Caillier SJ, Montalban X, Villoslada P, Fernandez O, Brassat D *et al.* Genome-wide pharmacogenomic analysis of the response to interferon beta therapy in multiple sclerosis. *Arch Neurol* 2008; **65**: 337-344.
182. Comabella M, Craig DW, Morcillo-Suarez C, Rio J, Navarro A, Fernandez M *et al.* Genome-wide scan of 500,000 single-nucleotide polymorphisms among responders and nonresponders to interferon beta therapy in multiple sclerosis. *Arch Neurol* 2009; **66**: 972-978.
183. Matute C, Domercq M, Sanchez-Gomez MV. Glutamate-mediated glial injury: mechanisms and clinical importance. *Glia* 2006; **53**: 212-224.

184. The expanding genetic overlap between multiple sclerosis and type I diabetes. *Genes Immun* 2009; **10**: 11-14.
185. Baker BJ, Akhtar LN, Benveniste EN. SOCS1 and SOCS3 in the control of CNS immunity. *Trends Immunol* 2009; **30**: 392-400.
186. Maier J, Kincaid C, Pagenstecher A, Campbell IL. Regulation of signal transducer and activator of transcription and suppressor of cytokine-signaling gene expression in the brain of mice with astrocyte-targeted production of interleukin-12 or experimental autoimmune encephalomyelitis. *Am J Pathol* 2002; **160**: 271-288.
187. Stark JL, Cross AH. Differential expression of suppressors of cytokine signaling-1 and -3 and related cytokines in central nervous system during remitting versus non-remitting forms of experimental autoimmune encephalomyelitis. *Int Immunol* 2006; **18**: 347-353.
188. Balabanov R, Strand K, Goswami R, McMahon E, Begolka W, Miller SD *et al.* Interferon-gamma-oligodendrocyte interactions in the regulation of experimental autoimmune encephalomyelitis. *J Neurosci* 2007; **27**: 2013-2024.
189. Mujtaba MG, Flowers LO, Patel CB, Patel RA, Haider MI, Johnson HM. Treatment of mice with the suppressor of cytokine signaling-1 mimetic peptide, tyrosine kinase inhibitor peptide, prevents development of the acute form of experimental allergic encephalomyelitis and induces stable remission in the chronic relapsing/remitting form. *J Immunol* 2005; **175**: 5077-5086.
190. Berard JL, Kerr BJ, Johnson HM, David S. Differential expression of SOCS1 in macrophages in relapsing-remitting and chronic EAE and its role in disease severity. *Glia* 2010; **58**: 1816-1826.
191. Tanaka K, Ichiyama K, Hashimoto M, Yoshida H, Takimoto T, Takaesu G *et al.* Loss of suppressor of cytokine signaling 1 in helper T cells leads to defective Th17 differentiation by enhancing antagonistic effects of IFN-gamma on STAT3 and Smads. *J Immunol* 2008; **180**: 3746-3756.
192. Gylvin T, Ek J, Nolsoe R, Albrechtsen A, Andersen G, Bergholdt R *et al.* Functional SOCS1 polymorphisms are associated with variation in obesity in whites. *Diabetes Obes Metab* 2009; **11**: 196-203.
193. Harada M, Nakashima K, Hirota T, Shimizu M, Doi S, Fujita K *et al.* Functional polymorphism in the suppressor of cytokine signaling 1 gene associated with adult asthma. *Am J Respir Cell Mol Biol* 2007; **36**: 491-496.
194. Zuvich RL, McCauley JL, Oksenberg JR, Sawcer SJ, De Jager PL, Aubin C *et al.* Genetic variation in the IL7RA/IL7 pathway increases multiple sclerosis susceptibility. *Hum Genet* 2010; **127**: 525-535.

195. Wykes SM, Nelson JE, Visscher DW, Djakiew D, Krawetz SA. Coordinate expression of the PRM1, PRM2, and TNP2 multigene locus in human testis. *DNA Cell Biol* 1995; **14**: 155-161.
196. Beyeen AD, Adzemovic MZ, Ockinger J, Stridh P, Becanovic K, Laaksonen H *et al.* IL-22RA2 associates with multiple sclerosis and macrophage effector mechanisms in experimental neuroinflammation. *J Immunol* 2010; **185**: 6883-6890.
197. Strange A, Capon F, Spencer CC, Knight J, Weale ME, Allen MH *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet* 2010; **42**: 985-990.
198. Lin PI, Vance JM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* 2007; **80**: 531-538.
199. Clarke GM, Cardon LR. Aspects of observing and claiming allele flips in association studies. *Genet Epidemiol* 2010; **34**: 266-274.
200. Burwick RM, Ramsay PP, Haines JL, Hauser SL, Oksenberg JR, Pericak-Vance MA *et al.* APOE epsilon variation in multiple sclerosis susceptibility and disease severity: some answers. *Neurology* 2006; **66**: 1373-1383.
201. Xuan C, Zhang BB, Li M, Deng KF, Yang T, Zhang XE. No association between APOE epsilon 4 allele and multiple sclerosis susceptibility: a meta-analysis from 5472 cases and 4727 controls. *J Neurol Sci* 2011; **308**: 110-116.
202. Rosche B, Cepok S, Stei S, Vogel F, Grummel V, Hoffmann S *et al.* The role of the polio virus receptor and the herpesvirus entry mediator B genes for the development of MS. *J Neuroimmunol* 2004; **156**: 171-177.
203. Santos M, Costa MC, Edite Rio M, Jose Sa M, Monteiro M, Valenca A *et al.* Genotypes at the APOE and SCA2 loci do not predict the course of multiple sclerosis in patients of Portuguese origin. *Mult Scler* 2004; **10**: 153-157.
204. ANZgene. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009; **41**: 824-828.
205. van der Walt A, Stankovich J, Bahlo M, Taylor BV, van der Mei IA, Foote SJ *et al.* Apolipoprotein genotype does not influence MS severity, cognition, or brain atrophy. *Neurology* 2009; **73**: 1018-1025.
206. International Multiple Sclerosis Genetics Consortium (IMSGC). Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. *Hum Mol Genet* 2010; **19**: 953-962.
207. Lessard CJ, Adrianto I, Ice JA, Wiley GB, Kelly JA, Glenn SB *et al.* Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci

- for systemic lupus erythematosus in a large-scale multiracial replication study. *Am J Hum Genet* 2012; **90**: 648-660.
208. Bergholdt R, Taxvig C, Eising S, Nerup J, Pociot F. CBLB variants in type 1 diabetes and their genetic interaction with CTLA4. *J Leukoc Biol* 2005; **77**: 579-585.
209. Payne F, Cooper JD, Walker NM, Lam AC, Smink LJ, Nutland S *et al.* Interaction analysis of the CBLB and CTLA4 genes in type 1 diabetes. *J Leukoc Biol* 2007; **81**: 581-583.
210. Marquez A, Mendoza JL, Taxonera C, Diaz-Rubio M, De La Concha EG, Urcelay E *et al.* IL23R and IL12B polymorphisms in Spanish IBD patients: no evidence of interaction. *Inflamm Bowel Dis* 2008; **14**: 1192-1196.
211. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011; **43**: 246-252.
212. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X *et al.* A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet* 2009; **41**: 1228-1233.
213. Capon F, Di Meglio P, Szaub J, Prescott NJ, Dunster C, Baumber L *et al.* Sequence variants in the genes for the interleukin-23 receptor (IL23R) and its ligand (IL12B) confer protection against psoriasis. *Hum Genet* 2007; **122**: 201-206.
214. Begovich AB, Chang M, Caillier SJ, Lew D, Catanese JJ, Wang J *et al.* The autoimmune disease-associated IL12B and IL23R polymorphisms in multiple sclerosis. *Hum Immunol* 2007; **68**: 934-937.
215. Alloza I, Heggarty S, Goris A, Graham CA, Dubois B, McDonnell G *et al.* Interleukin-12 p40 polymorphism and susceptibility to multiple sclerosis. *Ann Neurol* 2002; **52**: 524-525.
216. van Veen T, Crusius JB, Schrijver HM, Bouma G, Killestein J, van Winsen L *et al.* Interleukin-12p40 genotype plays a role in the susceptibility to multiple sclerosis. *Ann Neurol* 2001; **50**: 275.
217. Walsh JP, Berry J, Liu S, Panicker V, Dayan CM, Brix TH *et al.* The clinical presentation of autoimmune thyroid disease in men is associated with IL12B genotype. *Clin Endocrinol (Oxf)* 2011; **74**: 508-512.
218. Rossi S, Buttari F, Studer V, Motta C, Gravina P, Castelli M *et al.* The (AAT)_n repeat of the cannabinoid CB1 receptor gene influences disease progression in relapsing multiple sclerosis. *Mult Scler* 2011; **17**: 281-288.

219. Storr M, Emmerdinger D, Diegelmann J, Pfennig S, Ochsenkuhn T, Goke B *et al.* The cannabinoid 1 receptor (CNR1) 1359 G/A polymorphism modulates susceptibility to ulcerative colitis and the phenotype in Crohn's disease. *PLoS One* 2010; **5**: e9453.
220. Booth DR, Heard RN, Stewart GJ, Cox M, Scott RJ, Lechner-Scott J *et al.* Lack of support for association between the KIF1B rs10492972[C] variant and multiple sclerosis. *Nat Genet* 2010; **42**: 469-470; author reply 470-461.
221. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide BM *et al.* Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS Genet* 2012; **8**: e1002548.
222. Zhu S, Pan W, Song X, Liu Y, Shao X, Tang Y *et al.* The microRNA miR-23b suppresses IL-17-associated autoimmune inflammation by targeting TAB2, TAB3 and IKK-alpha. *Nat Med* 2012; **18**: 1077-1086.
223. Murugaiyan G, Beynon V, Mittal A, Joller N, Weiner HL. Silencing microRNA-155 ameliorates experimental autoimmune encephalomyelitis. *J Immunol* 2011; **187**: 2213-2221.
224. Junker A, Krumbholz M, Eisele S, Mohan H, Augstein F, Bittner R *et al.* MicroRNA profiling of multiple sclerosis lesions identifies modulators of the regulatory protein CD47. *Brain* 2009; **132**: 3342-3352.
225. Guan H, Fan D, Mrelashvili D, Hao H, Singh NP, Singh UP *et al.* MicroRNA let-7e is associated with the pathogenesis of experimental autoimmune encephalomyelitis. *Eur J Immunol* 2013; **43**: 104-114.
226. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA *et al.* A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061-1073.
227. Yang TP, Beazley C, Montgomery SB, Dimas AS, Gutierrez-Arcelus M, Stranger BE *et al.* Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* 2010; **26**: 2474-2476.
228. Zhang J, Salcedo TW, Wan X, Ullrich S, Hu B, Gregorio T *et al.* Modulation of T-cell responses to alloantigens by TR6/DcR3. *J Clin Invest* 2001; **107**: 1459-1468.
229. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; **489**: 57-74.

-
230. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**: 1790-1797.
231. Swaminathan B, Matesanz F, Cavanillas ML, Alloza I, Otaegui D, Olascoaga J *et al.* Validation of the CD6 and TNFRSF1A loci as risk factors for multiple sclerosis in Spain. *J Neuroimmunol* 2010; **223**: 100-103.
232. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B *et al.* The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225-2229.
233. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008; **24**: 2938-2939.
234. Vila J, Padilla O, Arman M, Gimferrer I, Lozano F. The scavenger receptor cysteine-rich superfamily (SCRC-SF). Structure and function of group B members. *Inmunología* 2000; **19**: 105-110.
235. Singer NG, Fox DA, Haqqi TM, Beretta L, Endres JS, Prohaska S *et al.* CD6: expression during development, apoptosis and selection of human and mouse thymocytes. *Int Immunol* 2002; **14**: 585-597.
236. Braun M, Muller B, ter Meer D, Raffegerst S, Simm B, Wilde S *et al.* The CD6 scavenger receptor is differentially expressed on a CD56 natural killer cell subpopulation and contributes to natural killer-derived cytokine and chemokine secretion. *J Innate Immun* 2011; **3**: 420-434.
237. Aruffo A, Melnick MB, Linsley PS, Seed B. The lymphocyte glycoprotein CD6 contains a repeated domain structure characteristic of a new family of cell surface and secreted proteins. *J Exp Med* 1991; **174**: 949-952.
238. Mayer B, Funke I, Seed B, Riethmuller G, Weiss E. Expression of the CD6 T lymphocyte differentiation antigen in normal human brain. *J Neuroimmunol* 1990; **29**: 193-202.
239. Kofler DM, Severson CA, Mousissian N, De Jager PL, Hafler DA. The CD6 multiple sclerosis susceptibility allele is associated with alterations in CD4+ T cell proliferation. *J Immunol* 2011; **187**: 3286-3291.
240. Heap GA, Yang JH, Downes K, Healy BC, Hunt KA, Bockett N *et al.* Genome-wide analysis of allelic expression imbalance in human primary cells by high-throughput transcriptome resequencing. *Hum Mol Genet* 2010; **19**: 122-134.
241. Castro MA, Oliveira MI, Nunes RJ, Fabre S, Barbosa R, Peixoto A *et al.* Extracellular isoforms of CD6 generated by alternative splicing regulate

- targeting of CD6 to the immunological synapse. *J Immunol* 2007; **178**: 4351-4361.
242. Bowen MA, Whitney GS, Neubauer M, Starling GC, Palmer D, Zhang J *et al*. Structure and chromosomal location of the human CD6 gene: detection of five human CD6 isoforms. *J Immunol* 1997; **158**: 1149-1156.
243. Cayrol R, Wosik K, Berard JL, Dodelet-Devillers A, Ifergan I, Kebir H *et al*. Activated leukocyte cell adhesion molecule promotes leukocyte trafficking into the central nervous system. *Nat Immunol* 2008; **9**: 137-145.
244. Gimferrer I, Calvo M, Mittelbrunn M, Farnos M, Sarrias MR, Enrich C *et al*. Relevance of CD6-mediated interactions in T cell activation and proliferation. *J Immunol* 2004; **173**: 2262-2270.
245. Oliveira MI, Goncalves CM, Pinto M, Fabre S, Santos AM, Lee SF *et al*. CD6 attenuates early and late signaling events, setting thresholds for T-cell activation. *Eur J Immunol* 2012; **42**: 195-205.
246. Gangemi RM, Swack JA, Gaviria DM, Romain PL. Anti-T12, an anti-CD6 monoclonal antibody, can activate human T lymphocytes. *J Immunol* 1989; **143**: 2439-2447.
247. Bott CM, Doshi JB, Morimoto C, Romain PL, Fox DA. Activation of human T cells through CD6: functional effects of a novel anti-CD6 monoclonal antibody and definition of four epitopes of the CD6 glycoprotein. *Int Immunol* 1993; **5**: 783-792.
248. Morimoto C, Rudd CE, Letvin NL, Hagan M, Schlossman SF. 2H1--a novel antigen involved in T lymphocyte triggering. *J Immunol* 1988; **140**: 2165-2170.
249. Hassan NJ, Barclay AN, Brown MH. Frontline: Optimal T cell activation requires the engagement of CD6 and CD166. *Eur J Immunol* 2004; **34**: 930-940.
250. Sarrias MR, Farnos M, Mota R, Sanchez-Barbero F, Ibanez A, Gimferrer I *et al*. CD6 binds to pathogen-associated molecular patterns and protects from LPS-induced septic shock. *Proc Natl Acad Sci U S A* 2007; **104**: 11724-11729.
251. Ibanez A, Sarrias MR, Farnos M, Gimferrer I, Serra-Pages C, Vives J *et al*. Mitogen-activated protein kinase pathway activation by the CD6 lymphocyte surface receptor. *J Immunol* 2006; **177**: 1152-1159.
252. Waite JC, Skokos D. Th17 response and inflammatory autoimmune diseases. *Int J Inflam* 2012; **2012**: 819467.
253. Curtis MM, Way SS. Interleukin-17 in host defence against bacterial, mycobacterial and fungal pathogens. *Immunology* 2009; **126**: 177-185.

-
254. Ke Y, Liu K, Huang GQ, Cui Y, Kaplan HJ, Shao H *et al.* Anti-inflammatory role of IL-17 in experimental autoimmune uveitis. *J Immunol* 2009; **182**: 3183-3190.
 255. O'Connor W, Jr., Kamanaka M, Booth CJ, Town T, Nakae S, Iwakura Y *et al.* A protective function for interleukin 17A in T cell-mediated intestinal inflammation. *Nat Immunol* 2009; **10**: 603-609.
 256. Durelli L, Conti L, Clerico M, Boselli D, Contessa G, Ripellino P *et al.* T-helper 17 cells expand in multiple sclerosis and are inhibited by interferon-beta. *Ann Neurol* 2009; **65**: 499-509.
 257. Kebir H, Kreyenborg K, Ifergan I, Dodelet-Devillers A, Cayrol R, Bernard M *et al.* Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. *Nat Med* 2007; **13**: 1173-1175.
 258. Du Y, Su Y, He J, Yang Y, Shi Y, Cui Y *et al.* Impact of the leucocyte immunoglobulin-like receptor A3 (LILRA3) on susceptibility and subphenotypes of systemic lupus erythematosus and Sjogren's syndrome. *Ann Rheum Dis* 2014.
 259. Kabalak G, Dobberstein SB, Matthias T, Reuter S, The YH, Dorner T *et al.* Association of immunoglobulin-like transcript 6 deficiency with Sjogren's syndrome. *Arthritis Rheum* 2009; **60**: 2923-2925.
 260. Hirayasu K, Ohashi J, Tanaka H, Kashiwase K, Ogawa A, Takanashi M *et al.* Evidence for natural selection on leukocyte immunoglobulin-like receptors for HLA class I in Northeast Asians. *Am J Hum Genet* 2008; **82**: 1075-1083.
 261. Wisniewski A, Luszczyk W, Manczak M, Jasek M, Kubicka W, Cislo M *et al.* Distribution of LILRA3 (ILT6/LIR4) deletion in psoriatic patients and healthy controls. *Hum Immunol* 2003; **64**: 458-461.
 262. Vandenbroeck K, Alloza I, Swaminathan B, Antiguedad A, Otaegui D, Olascoaga J *et al.* Validation of IRF5 as multiple sclerosis risk gene: putative role in interferon beta therapy and human herpes virus-6 infection. *Genes Immun* 2011; **12**: 40-45.
 263. Kristjansdottir G, Sandling JK, Bonetti A, Roos IM, Milani L, Wang C *et al.* Interferon regulatory factor 5 (IRF5) gene variants are associated with multiple sclerosis in three distinct populations. *J Med Genet* 2008; **45**: 362-369.
 264. Filer C, Ho P, Smith RL, Griffiths C, Young HS, Worthington J *et al.* Investigation of association of the IL12B and IL23R genes with psoriatic arthritis. *Arthritis Rheum* 2008; **58**: 3705-3709.

- 265. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E *et al.* Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 2014; **343**: 1246949.
- 266. Lee MN, Ye C, Villani AC, Raj T, Li W, Eisenhaure TM *et al.* Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* 2014; **343**: 1246980.
- 267. Kugelberg E. Immunogenetics: tracking immune activity across the genome. *Nat Rev Immunol* 2014; **14**: 212.
- 268. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet* 2013; **93**: 779-797.
- 269. Mkhikian H, Grigorian A, Li CF, Chen HL, Newton B, Zhou RW *et al.* Genetics and the environment converge to dysregulate N-glycosylation in multiple sclerosis. *Nat Commun* 2011; **2**: 334.
- 270. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* 2012; **337**: 1190-1195.
- 271. Jafari N, Broer L, van Duijn CM, Janssens AC, Hintzen RQ. Perspectives on the use of multiple sclerosis risk genes for prediction. *PLoS One* 2011; **6**: e26493.
- 272. Bahreini SA, Jabalameli MR, Saadatnia M, Zahednasab H. The role of non-HLA single nucleotide polymorphisms in multiple sclerosis susceptibility. *J Neuroimmunol* 2010; **229**: 5-15.

Anexo I: publicaciones

Replication study of 10 genes showing evidence for association with multiple sclerosis: validation of *TMEM39A*, *IL12B* and *CBLB* genes

Multiple Sclerosis Journal
18(7) 959–965
© The Author(s) 2012
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1352458511432741
msj.sagepub.com


Jezabel Varadé^{1*}, Manuel Comabella^{2*}, Miguel A Ortiz¹, Rafael Arroyo³, Oscar Fernández⁴, M Jesús Pinto-Medel⁵, María Fedetz⁶, Guillermo Izquierdo⁷, Miguel Lucas⁸, Carlos López Gómez⁴, Antonio Catalá Rabasa⁶, Antonio Alcina⁶, Fuencisla Matesanz^{6,7}, Iraide Alloza⁹, Alfredo Antigüedad¹⁰, María García-Barcina¹¹, David Otaegui¹², Javier Olascoaga¹³, Albert Saiz¹⁴, Yolanda Blanco¹⁴, Xavier Montalbán², Koen Vandenbroeck^{9,15*}, and Elena Urcelay^{1*}

Abstract

Background and objectives: Ten genes previously showing different evidence of association with multiple sclerosis have been selected to validate.

Methods: Eleven polymorphisms were genotyped with the iPLEX™ Sequenom in a well-powered collection of Spanish origin including 2863 multiple sclerosis cases and 2930 controls.

Results: Replication extended to the following polymorphisms: *PKN2* (rs305217), *GTF2B* (rs7538427), *EPHA4* (rs1517440), *YTHDF3* (rs12115114), *ANKFN1* (rs17758761) and *PTPRM* (rs4798571), which did not reach the threshold of significance in a follow-up of the first genome-wide association study (GWAS) conducted in multiple sclerosis; *TMEM39A* (rs1132200), which appeared as a newly identified susceptibility gene in the same study; a gene previously reaching GWAS significance in Italy, *CBLB* (rs9657904); *IL12B* (rs6887695, rs10045431), a susceptibility gene shared by diverse autoimmune diseases and, finally, another gene showing inconclusive association with multiple sclerosis, *CNR1* (rs1049353).

Conclusions: Pooled analysis corroborated the effect on MS predisposition of three genes: *TMEM39A* [rs1132200: $p_{M-H}=0.001$; OR_{M-H} (95% CI)= 0.84 (0.75–0.93)], *IL12B* [rs6887695: $p_{M-H}=0.03$; OR_{M-H} (95% CI)= 1.09 (1.01–1.17)] and *CBLB* [rs9657904: $p_{M-H}=0.01$; OR_{M-H} (95% CI)= 0.89 (0.81–0.97)].

Keywords

multiple sclerosis susceptibility, *TMEM39A*, *IL12B*, *CBLB*

Date received: 4th August 2011; revised: 31st October 2011; accepted: 17th November 2011

¹Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Spain.

²Centre d'Esclerosi Múltiple de Catalunya, CEM-Cat, Unitat de Neuroimmunologia Clínica, Hospital Universitari Vall d'Hebron (HUVH), Barcelona, Spain.

³Multiple Sclerosis Unit, Neurology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Spain.

⁴Servicio de Neurología, Instituto de Neurociencias Clínicas, Hospital Regional Universitario Carlos Haya, Málaga, Spain.

⁵Laboratorio de Investigación, Instituto de Neurociencias Clínicas, Hospital Regional Universitario Carlos Haya, Málaga, Spain.

⁶Instituto Parasitología y Biomedicina 'López Neyra', C. S. I. C., Granada, Spain.

⁷Unidad de Esclerosi Múltiple, Hospital Virgen Macarena, Seville, Spain.

⁸Servicio de Biología Molecular, Hospital Virgen Macarena, Seville, Spain.

⁹Neurogenomiks Group, Universidad del País Vasco (UPV/EHU), Leioa, Spain.

¹⁰Servicio de Neurología, Hospital de Basurto, Spain.

¹¹Servicio de Genética, Hospital de Basurto, Spain.

¹²Área de Neurociencias, Instituto Investigación Sanitaria Biodonostia, San Sebastian, Spain.

¹³Servicio de Neurología, Unidad de Esclerosi Múltiple, Hospital Donostia, San Sebastian, Spain.

¹⁴Neurology Service, Hospital Clinic and Instituto d'Investigació Biomèdica Pi i Sunyer (IDIBAPS), Barcelona, Spain.

¹⁵IKERBASQUE, Basque Foundation for Science, Spain.

*These authors contributed equally to this work.

Corresponding author:

Elena Urcelay, PhD, Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC) Martin Lagos s/n, 28040 Madrid, Spain.
Email: eurcelay.hcsc@salud.madrid.org

ORIGINAL ARTICLE

A cytokine gene screen uncovers SOCS1 as genetic risk factor for multiple sclerosis

K Vandenbroeck^{1,2}, J Alvarez³, B Swaminathan¹, I Alloza¹, F Matesanz⁴, E Urcelay⁵, M Comabella⁶, A Alcina⁴, M Fedetz⁴, MA Ortiz⁵, G Izquierdo⁷, O Fernandez⁸, N Rodriguez-Ezpeleta³, C Matute⁹, S Caillier¹⁰, R Arroyo⁵, X Montalban⁶, JR Oksenberg¹⁰, A Antigüedad¹¹ and A Aransay³

¹Neurogenomiks Laboratory, Department of Neuroscience, University of the Basque Country UPV/EHU, Leioa, Spain; ²IKERBASQUE, Basque Foundation for Science, Bilbao, Spain; ³CIC bioGUNE, Parque Tecnológico de Bizkaia, Derio, Spain; ⁴Instituto de Parasitología y Biomedicina 'López Neyra', Consejo Superior de Investigaciones Científicas, Granada, Spain; ⁵Immunology and Neurology Department, Hospital Clínico S Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain; ⁶Centre d'Esclerosi Múltiple de Catalunya, CEM-Cat, Unitat de Neuroimmunologia Clínica, Hospital Universitari Vall d'Hebron, Barcelona, Spain; ⁷Unidad de Esclerosis Múltiple, Hospital Virgen Macarena, Sevilla, Spain; ⁸Servicio de Neurología, Instituto de Neurociencias Clínicas del Hospital Regional Universitario Carlos Haya de Málaga, Málaga, Spain; ⁹Neurotek Laboratory, Department of Neuroscience, University of the Basque Country UPV/EHU, Leioa, Spain; ¹⁰Department of Neurology, University of California, San Francisco, CA, USA and ¹¹Servicio de Neurología, Hospital de Basurto, Bilbao, Spain

Cytokine and cytokine receptor genes, including *IL2RA*, *IL7R* and *IL12A*, are known risk factors for multiple sclerosis (MS). Excitotoxic oligodendroglial death mediated by glutamate receptors contributes to demyelinating reactions. In the present study, we screened 368 single-nucleotide polymorphisms (SNPs) in 55 genes or gene clusters coding for cytokines, cytokine receptors, suppressors of cytokine signaling (SOCS), complement factors and glutamate receptors for association with MS in a Spanish–Basque resident population. Top-scoring SNPs were found within or nearby the genes coding for *SOCS1* ($P = 0.0005$), interleukin-28 receptor, alpha chain ($P = 0.0008$), oncostatin M receptor ($P = 0.002$) and interleukin-22 receptor, alpha 2 (*IL22RA2*; $P = 0.003$). The *SOCS1* rs243324 variant was validated as risk factor for MS in a separate cohort of 3919 MS patients and 4003 controls (combined Cochran–Mantel–Haenszel $P = 0.00006$; odds ratio (OR) = 1.13; 95% confidence interval (CI) = 1.07–1.20). In addition, the T allele of rs243324 was consistently increased in relapsing–remitting/secondary progressive versus primary–progressive MS patients, in each of the six data sets used in this study ($P_{CMH} = 0.0096$; OR = 1.24; 95% CI 1.05–1.46). The association with *SOCS1* appears independent from the chr16MS risk locus *CLEC16A*.

Genes and Immunity (2012) 13, 21–28; doi:10.1038/gene.2011.44; published online 30 June 2011

Keywords: multiple sclerosis; *SOCS1*; cytokine; genetics; single-nucleotide polymorphism

Introduction

Multiple sclerosis (MS) is a chronic (CH) inflammatory demyelinating disorder of the central nervous system of unknown etiology, and represents the most common cause of non-traumatic neurological disability in young adults. The observed rates of familial aggregation of MS reflected in the increased risk of siblings, second- and third-degree relatives to develop the disease, as well as twin studies, collectively reject a Mendelian trait as the driving force for susceptibility, but are reconcilable with a polygenic, multifactorial mechanism.¹ Although the human leukocyte antigen (HLA) gene cluster on chromosome 6p21.3 has

been known as susceptibility locus since the early 1970s, it was not until 2007, when the first non-HLA genetic risk factors were unequivocally identified via a genome-wide association study (GWAS).² Subsequently, seven more GWAS have led to the identification of around 15 validated non-HLA risk loci for MS, including among others *IL2RA*, *IL7R*, *CD58*, *EVI5* and *CD40*.¹ In addition, a meta-analysis of GWAS³ identified the additional loci *CD6*, *TNFRSF1A* and *IRF8*, and three further loci with suggestive evidence arising from this study were subsequently validated as genuine MS risk factors, that is, *IL12A*, *MPHOSPH9* and *RSG1*.⁴ All non-HLA MS-susceptibility alleles known so far are relatively common in the population and contribute only modestly to overall risk (odds ratios (OR) of 1.1–1.3).

In the present study, we report the results of a haptag screen primarily focusing on cytokine, cytokine receptor genes and associated signal transduction factors that also covered a small selection of ionotropic glutamate receptors and transporters. The latter category of genes was included based on the observation that glutamate-

Correspondence: Dr K Vandenbroeck, Neurogenomiks Laboratory, Department of Neuroscience, Universidad del País Vasco (UPV/EHU), Edificio 205, Planta-1, Parque Tecnológico de Bizkaia, 48170 Zamudio (Bizkaia), Spain.
E-mail: k.vandenbroeck@ikerbasque.org
Received 17 March 2011; revised 5 May 2011; accepted 11 May 2011; published online 30 June 2011

SHORT REPORT

Closing the case of *APOE* in multiple sclerosis: no association with disease risk in over 29 000 subjects

Christina M Lill,^{1,2} Tian Liu,^{2,3} Brit-Maren M Schjeide,² Johannes T Roehr,² Denis A Akkad,⁴ Vincent Damotte,⁵ Antonio Alcina,⁶ Miguel A Ortiz,⁷ Rafa Arroyo,⁸ Aitzkoa Lopez de Lapuente,⁹ Paul Blaschke,¹⁰ Alexander Winkelmann,¹⁰ Lisa-Ann Gerdes,¹¹ Felix Luessi,¹ Oscar Fernandez,¹² Guillermo Izquierdo,¹² Alfredo Antigüedad,¹³ Sabine Hoffjan,⁴ Isabelle Cournu-Rebeix,^{5,14} Silvana Gromöller,² Hans Faber,¹⁵ Maria Liebsch,² Esther Meissner,² Coralie Chanvillard,¹⁶ Emmanuel Touze,¹⁷ Fernando Pico,¹⁸ Philippe Corcia,¹⁹ ANZgene Consortium,[†] Thomas Dörner,²⁰ Elisabeth Steinhagen-Thiessen,²¹ Lars Baeckman,²² Hauke R Heekeren,^{3,23} Shu-Chen Li,³ Ulman Lindenberger,³ Andrew Chan,²⁴ Hans-Peter Hartung,²⁵ Orhan Aktas,²⁵ Peter Lohse,²⁶ Tania Kümpfel,¹¹ Christian Kubisch,²⁷ Joerg T Epplen,⁴ Uwe K Zettl,¹⁰ Bertrand Fontaine,^{5,12} Koen Vandenbroeck,^{9,28} Fuencisla Matesanz,⁶ Elena Urcelay,⁸ Lars Bertram,² Frauke Zipp,¹

► Additional supplementary files are published online only. To view these files please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2012-101175>).

For numbered affiliations see end of article

Correspondence to

Professor Frauke Zipp, Rhine Main Neuroscience Network (rmn2), Johannes Gutenberg University Center Mainz, Neurology Department, Langenbeckstr. 1, Mainz 55131, Germany; frauke.zipp@unimedizin-mainz.de

†A list of all co-authors for the ANZgene consortium can be found at the end of this manuscript

LB and FZ contributed equally

Received 17 July 2012
Revised 23 July 2012
Accepted 23 July 2012

ABSTRACT

Background Single nucleotide polymorphisms (SNPs) rs429358 (ε4) and rs7412 (ε2), both invoking changes in the amino-acid sequence of the apolipoprotein E (*APOE*) gene, have previously been tested for association with multiple sclerosis (MS) risk. However, none of these studies was sufficiently powered to detect modest effect sizes at acceptable type-I error rates. As both SNPs are only imperfectly captured on commonly used microarray genotyping platforms, their evaluation in the context of genome-wide association studies has been hindered until recently.

Methods We genotyped 12 740 subjects hitherto not studied for their *APOE* status, imputed raw genotype data from 8739 subjects from five independent genome-wide association studies datasets using the most recent high-resolution reference panels, and extracted genotype data for 8265 subjects from previous candidate gene assessments.

Results Despite sufficient power to detect associations at genome-wide significance thresholds across a range of ORs, our analyses did not support a role of rs429358 or rs7412 on MS susceptibility. This included meta-analyses of the combined data across 13 913 MS cases and 15 831 controls (OR=0.95, p=0.259, and OR 1.07, p=0.0569, for rs429358 and rs7412, respectively).

Conclusion Given the large sample size of our analyses, it is unlikely that the two *APOE* missense SNPs studied here exert any relevant effects on MS susceptibility.

INTRODUCTION

Multiple sclerosis (MS), the most common chronic inflammatory disease of the central nervous system, is likely caused by interplay of environmental and genetic factors. Recent genome-wide

association studies (GWAS) have identified almost 60 putative risk loci apart from a well-established association with the human leukocyte antigen region on chromosome 6p21 (eg, refs.^{1–3}). Despite this recent progress, it has been estimated that approximately 80% or more of the genetic variance remains unexplained by the currently known loci.³ The failure to decipher the full genetic spectrum of MS susceptibility may in part be due to incomplete coverage of the genome by the available GWAS arrays. Along these lines, two single nucleotide polymorphisms (SNPs) frequently tested for association with MS risk in the pre-GWAS era, that is, rs429358 (a.k.a. 'ε4') and rs7412 ('ε2') in encoding apolipoprotein E (*APOE*), are absent from most GWAS genotyping platforms. Thus, these two *APOE* SNPs could not be directly assessed in previous MS GWAS and GWAS meta-analyses. The most recent release of whole-genome sequence data from the 1000 Genomes Project now makes it possible to impute genotypes at both sites for data from most GWAS genotyping platforms.

The investigation of *APOE*, the single most important risk locus for Alzheimer's disease, in MS has been motivated by reports of genetic linkage to the *APOE*-containing region on chromosome 19q13 (eg, ref.⁴) as well as by its functional role in lipid transport, immunoregulation, neuroplasticity and repair mechanisms.⁵ However, *APOE* association studies in MS have yielded mostly negative results to date with some studies reporting significant effects while most others were unable to confirm these associations. These inconsistencies can at least in part be attributed to small sample sizes of the individual datasets. As a matter of fact, none of the previously performed *APOE* association studies in MS included more

MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis

International Multiple Sclerosis Genetics Consortium[#]

[#]A list of all authors and their affiliations can be found in Appendix 1;

Correspondence to: Prof. Dr. Frauke Zipp,
Rhine Main Neuroscience Network (rmn2),
Johannes Gutenberg University Centre Mainz,
Neurology Department,
Langenbeckstr. 1,
55131 Mainz, Germany,
E-mail: frauke.zipp@unimedizin-mainz.de

Correspondence may also be addressed to: Dr. Christina M. Lill, Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Neuropsychiatric Genetics Group, Ihnestr. 63-73, 14195 Berlin, Germany, E-mail: lill@molgen.mpg.de

A recent genome-wide association study reported five loci for which there was strong, but sub-genome-wide significant evidence for association with multiple sclerosis risk. The aim of this study was to evaluate the role of these potential risk loci in a large and independent data set of ~20 000 subjects. We tested five single nucleotide polymorphisms rs228614 (MANBA), rs630923 (CXCR5), rs2744148 (SOX8), rs180515 (RPS6KB1), and rs6062314 (ZBTB46) for association with multiple sclerosis risk in a total of 8499 cases with multiple sclerosis, 8765 unrelated control subjects and 958 trios of European descent. In addition, we assessed the overall evidence for association by combining these newly generated data with the results from the original genome-wide association study by meta-analysis. All five tested single nucleotide polymorphisms showed consistent and statistically significant evidence for association with multiple sclerosis in our validation data sets (rs228614: odds ratio = 0.91, $P = 2.4 \times 10^{-6}$; rs630923: odds ratio = 0.89, $P = 1.2 \times 10^{-4}$; rs2744148: odds ratio = 1.14, $P = 1.8 \times 10^{-6}$; rs180515: odds ratio = 1.12, $P = 5.2 \times 10^{-7}$; rs6062314: odds ratio = 0.90, $P = 4.3 \times 10^{-3}$). Combining our data with results from the previous genome-wide association study by meta-analysis, the evidence for association was strengthened further, surpassing the threshold for genome-wide significance ($P < 5 \times 10^{-8}$) in each case. Our study provides compelling evidence that these five loci are genuine multiple sclerosis susceptibility loci. These results may eventually lead to a better understanding of the underlying disease pathophysiology.

Keywords: multiple sclerosis; complex genetics; genetic risk; immunogenetics; genetic association

Abbreviations: GWAS = genome-wide association study; SNP = single nucleotide polymorphism

Introduction

Multiple sclerosis is the most common inflammatory demyelinating disease of the CNS that is likely caused by an interplay of genetic and environmental risk factors. Apart from several independent association signals in the *HLA* (human leukocyte antigen) region

on chromosome 6p21, a recent genome-wide association study (GWAS) in multiple sclerosis has reported 52 loci exerting small to moderate risk effects (IMSGC and WTCCC2, 2011). In addition, five additional loci provided strong support for association ($P < 5 \times 10^{-7}$) in that GWAS, but failed to meet current criteria for genome-wide significance ($P < 5 \times 10^{-8}$). The most

Fine Mapping and Functional Analysis of the Multiple Sclerosis Risk Gene CD6

Bhairavi Swaminathan¹, Angélica Cuapio², Iraide Alloza¹, Fuencisla Matesanz³, Antonio Alcina³, Maria García-Barcina⁴, Maria Fedetz³, Óscar Fernández⁵, Miguel Lucas⁶, Teresa Órpez⁷, M^a Jesus Pinto-Medel⁷, David Otaegui⁸, Javier Olascoaga⁹, Elena Urcelay¹⁰, Miguel A. Ortiz¹⁰, Rafael Arroyo¹¹, Jorge R. Oksenberg¹², Alfredo Antigüedad¹³, Eva Tolosa², Koen Vandenbroeck^{1,14*}

1 Neurogenomiks Laboratory, University of the Basque Country (UPV/EHU), Leioa, Spain, **2** Department of Immunology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, **3** Instituto de Parasitología y Biomedicina "López Neyra" Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain, **4** Servicio de Genética, Hospital de Basurto, Bilbao, Spain, **5** Department of Neurology, Institute of Clinical Neurosciences, Hospital Regional Universitario Carlos Haya, Málaga, Spain, **6** Unidad de Esclerosis Múltiple, Hospital Virgen Macarena, Sevilla, Spain, **7** Research Laboratory, Institute of Clinical Neurosciences, Hospital Regional Universitario Carlos Haya, Málaga, Spain, **8** Área de Neurociencias, Instituto de Investigación Sanitaria Biodonostia, San Sebastián, Spain, **9** Servicio de Neurología, Unidad de Esclerosis Múltiple, Hospital Donostia, San Sebastián, Spain, **10** Immunology Department H. Clínico S. Carlos, Instituto de Investigación Sanitaria S. Carlos (IdISSC), Madrid, Spain, **11** Multiple Sclerosis Unit, Neurology Department H. Clínico S. Carlos, Instituto de Investigación Sanitaria S. Carlos (IdISSC), Madrid, Spain, **12** Department of Neurology, University of California San Francisco, San Francisco, California, United States of America, **13** Servicio de Neurología, Hospital de Basurto, Bilbao, Spain, **14** IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Abstract

CD6 has recently been identified and validated as risk gene for multiple sclerosis (MS), based on the association of a single nucleotide polymorphism (SNP), rs17824933, located in intron 1. CD6 is a cell surface scavenger receptor involved in T-cell activation and proliferation, as well as in thymocyte differentiation. In this study, we performed a haptag SNP screen of the CD6 gene locus using a total of thirteen tagging SNPs, of which three were non-synonymous SNPs, and replicated the recently reported GWAS SNP rs650258 in a Spanish-Basque collection of 814 controls and 823 cases. Validation of the six most strongly associated SNPs was performed in an independent collection of 2265 MS patients and 2600 healthy controls. We identified association of haplotypes composed of two non-synonymous SNPs [rs11230563 (R225W) and rs2074225 (A257V)] in the 2nd SRCR domain with susceptibility to MS ($P_{\text{max(T) permutation}} = 1 \times 10^{-4}$). The effect of these haplotypes on CD6 surface expression and cytokine secretion was also tested. The analysis showed significantly different CD6 expression patterns in the distinct cell subsets, i.e. – CD4⁺ naïve cells, $P=0.0001$; CD8⁺ naïve cells, $P<0.0001$; CD4⁺ and CD8⁺ central memory cells, $P=0.01$ and 0.05 , respectively; and natural killer T (NKT) cells, $P=0.02$; with the protective haplotype (RA) showing higher expression of CD6. However, no significant changes were observed in natural killer (NK) cells, effector memory and terminally differentiated effector memory T cells. Our findings reveal that this new MS-associated CD6 risk haplotype significantly modifies expression of CD6 on CD4⁺ and CD8⁺ T cells.

Citation: Swaminathan B, Cuapio A, Alloza I, Matesanz F, Alcina A, et al. (2013) Fine Mapping and Functional Analysis of the Multiple Sclerosis Risk Gene CD6. PLoS ONE 8(4): e62376. doi:10.1371/journal.pone.0062376

Editor: Pablo Villoslada, Institute Biomedical Research August Pi Sunyer (IDIBAPS) - Hospital Clinic of Barcelona, Spain

Received: March 20, 2012; **Accepted:** March 22, 2013; **Published:** April 24, 2013

Copyright: © 2013 Swaminathan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by grants to K.V. from the European Community's Seventh Framework Programme [FP7/2007–2013] under grant agreement no. 212877 (UEPHA*MS; www.reem.es/uepha-ms/) and from the Gobierno Vasco (Grupos de Investigación del Sistema Universitario Vasco; ref. ITS12-10). B.S. and A.C. are early-stage researchers of UEPHA*MS (No 2121877). Grants to A.A. and F.M. were provided by Ministerio de Ciencia e Innovación - FEDER (SAF2009-11491) and FIS_FEDER (CP10/00526), Junta de Andalucía-FEDER (P07-CVI-02551). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: k.vandenbroeck@ikerbasque.org

Introduction

Multiple sclerosis (MS) is a disorder of the central nervous system that is characterized by chronic inflammation, demyelination, axonal loss and neurodegeneration [1]. Genome-wide association (GWAS) screens and meta-analyses have enabled identification of about 50 non-HLA MS risk genes [2–16]. Apart from the HLA region, the implicated genes exert modest effects at the population level with odds ratios (OR) ranging from 1.1–1.3 [17]. In a recent study, we validated the association of four risk single nucleotide polymorphisms (SNPs) with MS susceptibility [5], in a northern Spanish-Basque population, from which the CD6 SNP rs17824933 emerged with a stronger risk (OR = 1.34) [18].

CD6 is a member of the group B scavenger receptor cysteine-rich super family (SRCR-SF) [19] found on thymocytes, mature T-cells, some B-cell and natural killer (NK) subsets and is also expressed in certain parts of the brain like the cerebellum, basal ganglia, thalamus, corpus amygdaloideum, and cerebral lobi [20–23].

At the transcriptional level, in addition to the full-length form, a total of six isoforms have been reported that diversify the cytoplasmic domains [24,25]. The ligand for CD6 is the Activated Leukocyte Cell Adhesion Molecule (ALCAM), found in the thymic epithelium and in the epithelial layer of the blood-brain barrier. ALCAM binds to the third SRCR domain of CD6 and this

Influence of the *LILRA3* deletion on the risk of developing multiple sclerosis: a meta-analysis

Miguel A. Ortiz¹, Concepción Núñez¹, David Ordóñez², José C. Alvarez-Cermeño³, José E. Martínez-Rodríguez⁴, Antonio J. Sánchez⁵, Rafael Arroyo⁶, Guillermo Izquierdo⁷, Sunny Malhotra⁸, Xavier Montalban⁸, Antonio García-Merino⁵, Elvira Munteis⁴, Antonio Alcina⁹, Manuel Comabella⁸, Fuencisla Matesanz⁹, Luisa M. Villar^{*3}, Elena Urcelay^{*1}

¹ Immunology Department. Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Spain.

² Immunogenetics & Histocompatibility, Instituto de Investigación Sanitaria Puerta de Hierro, Majadahonda, Madrid, Spain.

³ Departments of Immunology and Neurology, Multiple Sclerosis Unit, Hospital Ramon y Cajal, (IRYCIS), Madrid, Spain.

⁴ Neurology, Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain

⁵ Neuroimmunology, Hospital Universitario Puerta de Hierro, Majadahonda, Madrid, Spain

⁶ Multiple Sclerosis Unit, Neurology Department. Hospital Clínico S. Carlos, Instituto de Investigación Sanitaria S. Carlos (IdISSC), Madrid, Spain.

⁷ Multiple Sclerosis Unit, Hospital Virgen Macarena, Sevilla, Spain.

⁸ Servei de Neurologia-Neuroimmunologia, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain.

⁹ Department of Cell Biology and Immunology, Instituto de Parasitología y Biomedicina “López Neyra”, Consejo Superior de Investigaciones Científicas (IPBLN-CSIC), Granada, Spain.

*These authors contributed equally to this work.

Running title: *LILRA3* deletion on multiple sclerosis risk

Corresponding Author:

Elena Urcelay, PhD.
Immunology Department,
Hospital Universitario Clínico San Carlos,
Martin Lagos s/n, 28040 Madrid, Spain.
Phone: 34- 91 330 3000 7171
e-mail: elena.urcelay@salud.madrid.org

Abstract

Multiple sclerosis (MS) is a neurodegenerative, autoimmune disease of the central nervous system. Genome-wide association studies (GWAS) have identified over hundred polymorphisms with modest individual effects in MS susceptibility and have confirmed that the Major Histocompatibility Complex exerts the greatest individual effect. Additional risk loci with immunologically relevant genes were found significantly overrepresented. Nonetheless, it is accepted that most of the genetic architecture underlying susceptibility to the disease remains to be defined. Candidate association studies of the leukocyte immunoglobulin-like receptor *LILRA3* gene in MS have been repeatedly reported with inconsistent results. In an attempt to shed some light on these controversial findings, a combined analysis was performed including the previously published datasets and three newly genotyped cohorts. Both wild-type and deleted *LILRA3* alleles were discriminated in a single-tube PCR amplification and the resulting products were visualized by their different electrophoretic mobilities. Overall, this meta-analysis involved 3200 MS patients and 3069 matched healthy controls and did not evidence significant association of the *LILRA3* deletion [carriers of *LILRA3* deletion: $p=0.25$, OR (95% CI)= 1.07 (0.95-1.19)], even after stratification by gender and the *HLA-DRB1*15:01* risk allele. In conclusion, we were unable to find any evidence supporting the putative association of the *LILRA3* gene with MS risk.

Key words: *LILRA3* deletion, multiple sclerosis, susceptibility

Introduction

Multiple sclerosis (MS), a disease of the central nervous system involving chronic inflammation, axonal damage and demyelination, is likely caused by the interplay of genetic and environmental risk factors. The class II *HLA-DRB1*15:01* allele has been described as the main genetic susceptibility factor for the disease [1], although other immunological relevant genes have been also recently associated with MS predisposition [2].

Leukocyte immunoglobulin-like receptors (LILRs), also known as immunoglobulin-like transcripts (ITLs) or monocyte inhibitory receptors (MIRs), are broadly expressed on myeloid and lymphoid cells and play an important role in modulating innate and adaptive immune responses [3]. LILRA3 is a soluble protein of the LILR family that contains four Ig-like domains (D1, D2, D3 and D4) and acts as a soluble receptor for certain class I Major Histocompatibility Complex (MHC) antigens [4].

Members of the *LILR* gene family map to the leukocyte receptor complex on chromosome 19q13.4, a region that encodes at least 24 members of the immunoglobulin superfamily [5]. Reportedly, chromosome 19q13 is one of the genomic regions that exhibits clustering of susceptibility loci in several autoimmune diseases including MS [6, 7]. However, genome-wide association studies (GWAS) performed to date have not been able to detect any signal with significant association in this region when considering MS risk [8, 9]. GWAS benefit of a high throughput technology based on genotyping a growing number of single-nucleotide polymorphisms (SNPs), even over a million. Nonetheless, this is a limited strategy to detect any other kind of genetic variants such as deletions, unless they show high linkage disequilibrium with a SNP. An unprecedented advance unraveling the genetic architecture of complex diseases has been driven by these genome-wide

studies, but a high percentage of the heritability of these disorders still remains unidentified [10].

The *LILRA3* gene exhibits a 6.7kb deletion comprising the first 6 of a total of 7 exons, resulting in a “null allele” because it removes all the regions coding Ig-like domains of the protein [11-13]. At present, little is known about its role in the immune system. The *LILRA3* gene has been associated with increased susceptibility to Sjögren's syndrome and to some subphenotypes of systemic lupus erythematosus [14, 15]. Increased levels of the soluble *LILRA3* protein have been found in patients with rheumatoid arthritis [16]. Moreover, a study published by Du *et al.* [17] provided evidence for the functional *LILRA3* allele as a genetic risk factor only for male patients with rheumatoid arthritis. These data point to a potential genetic background shared by different autoimmune diseases, as described by some authors [18].

The association of the *LILRA3* deletion with MS susceptibility has been repeatedly studied [6, 19-21] yielding inconclusive results, some of them reporting significant effects while others being unable to confirm the association. Therefore, we aimed to perform a new analysis trying to ascertain the effect of the *LILRA3* deletion on MS risk in people with European ancestry. To do so, all previously published datasets and three newly genotyped cohorts were included in a meta-analysis. Further, we investigated the epistatic interaction between *LILRA3* and the classical MS risk factor *HLA-DRB1*15:01*, considering also the effect of gender previously described as a determinant in the *LILRA3* effect.

Material and Methods

Patients

A total of 1932 relapsing-remitting and secondary progressive (11% of total) MS patients of Spanish ancestry and 1599 ethnically matched controls was consecutively recruited from the following Spanish Hospitals: H. Clínico San Carlos and H. Ramón y Cajal, from Madrid, Central Spain (1422 patients and 841 controls); H. Virgen Macarena (Sevilla) and H. Virgen de las Nieves (Granada) from Andalusia, South of Spain (236 patients and 443 controls); and H. Vall d'Hebron (Catalonia), Northeast of Spain (274 patients and 315 controls). Both clinical forms are considered two stages of the same disease, to which we will refer here as relapse-onset MS (R-MS). Patients were diagnosed on the basis of McDonald criteria [22] and were recruited after written informed consent. The Ethics Committees of the participant hospitals approved this study.

Genotyping

The presence/ absence of the *LILRA3* deletion in our samples was studied as previously described [20]. Briefly, both wild-type and deleted *LILRA3* alleles were discriminated in a single-tube PCR amplification and the resulting products were visualized by their different electrophoretic mobility in regular agarose gels.

Literature search and data abstraction

Genetic association studies investigating the role of *LILRA3* on MS susceptibility in European populations were identified by a search of the following terms: “*LILRA3* or *ITL6* or 19q13” and “multiple sclerosis” in NCBI's PubMed database and by inspecting

cross-references in related publications. All the articles published before September 2014 were included. Raw data were requested to the corresponding author when not included in the study.

Meta-analysis

Meta-analyses were performed with the published and the newly generated datasets, Fig. 1. The Cochran-Mantel-Haenszel method implemented in Review Manager (RevMan) version 5.0 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008) was used to calculate combined odds ratios (OR) and 95% confidence intervals (CI). P-values <0.05 were considered statistically significant. The Der Simonian and Laird random effects model was used according to the results of the tests of heterogeneity. A sensitivity analysis was performed to test the relative influence of each study: studies were sequentially dropped, and the effect on the change in the overall degree of heterogeneity was quantified with the I^2 statistic.

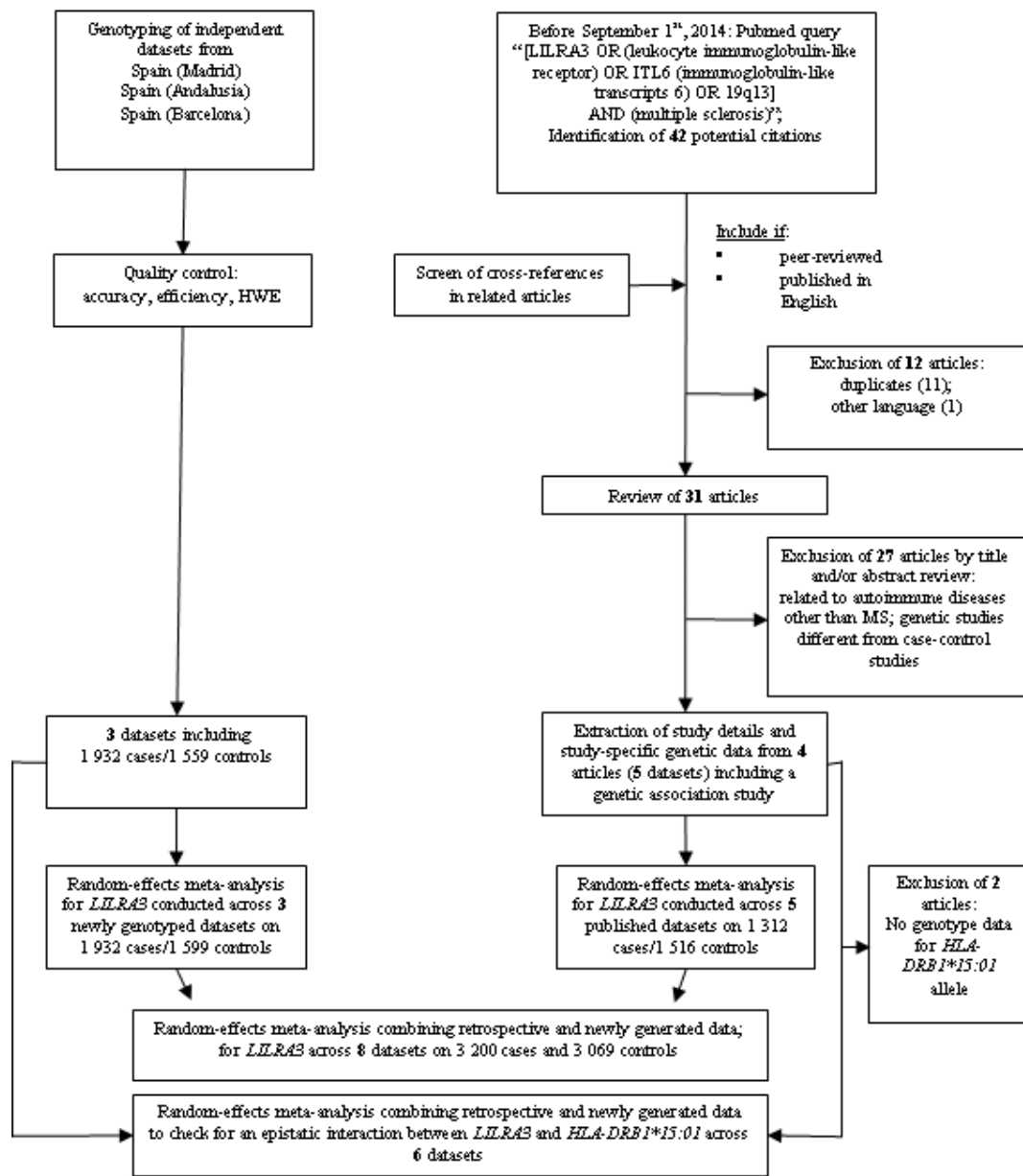


Fig. 1. Flowchart of literature search, data extraction and association analysis of *LILRA3* and MS risk.

Statistical analysis

Hardy-Weinberg proportions (HWE) were tested in controls. Statistical analyses were performed by using the SPSS 17.0 package (SPSS Inc, Chicago, IL). Allelic and genotypic frequency comparisons were analyzed using the χ^2 test of a Fisher exact test (when expected values were below 5). Odds ratio (ORs) were calculated and their 95% CIs were estimated with the Cornfield method.

Results

Association of the *LILRA3* deletion with R-MS susceptibility

Four previously published reports have studied the association of the *LILRA3* deletion with MS susceptibility risk in five individual datasets. Meta-analysis across these published datasets yielded a statistically significant effect in carriers of the *LILRA3* deletion [$p = 0.03$; OR (95% CI) = 1.19 (1.02-1.40)], with modest evidence of heterogeneity ($I^2 = 25\%$). Subsequently, a combined analysis of the previously published datasets adding the data of the three newly genotyped independent cohorts was performed. Overall, this meta-analysis including 3200 MS patients and 3069 matched healthy controls did not reveal significant evidence for association and yielded an effect size close to one [Fig. 2, carriers of *LILRA3* deletion: $p = 0.25$, OR (95% CI) = 1.07 (0.95-1.19)] with moderate evidence for heterogeneity ($I^2 = 32\%$). A further sensitivity analysis was performed in this study, which provided I^2 values ranging from 2% to 41%: the analysis excluding the dataset “2009-Spain (Madrid)” led to the lowest heterogeneity, and again statistical significance was not observed ($p = 0.49$).

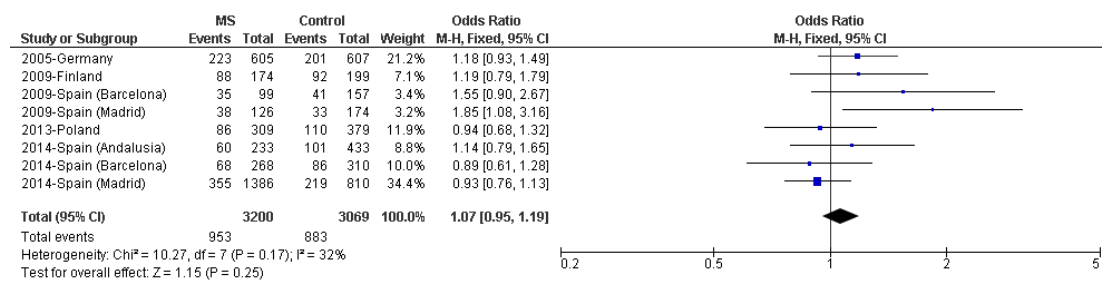


Fig. 2. Meta-analysis assessing the association of *LILRA3* deletion with MS risk in populations of European descent. Forest plot of random effects with study-specific ORs (squares) and 95% CIs (lines) calculated for each individual dataset. Pooled OR (diamond) and 95% CI were calculated combining all datasets.

Analysis of the epistatic interaction between the *LILRA3* deletion and the *HLA-DRB1*15:01* allele

To check for an epistatic interaction between the *LILRA3* deletion and the classical MS risk factor *HLA-DRB1*15:01*, a meta-analysis of those datasets including both variants was performed (Fig. 3). The meta-analysis did not evidence any significant epistatic interaction when these two genetic factors were considered [$p = 0.36$, OR (95% CI) = 1.09 (0.91-1.30)]. Conversely, to ascertain whether a possible effect of the *LILRA3* deletion might be restricted to subjects lacking the *HLA-DRB1*15:01* allele, the same analysis was performed in this group of patients, but a significant difference was not observed [$p = 0.74$, OR (95% CI) = 1.03 (0.88-1.20)].

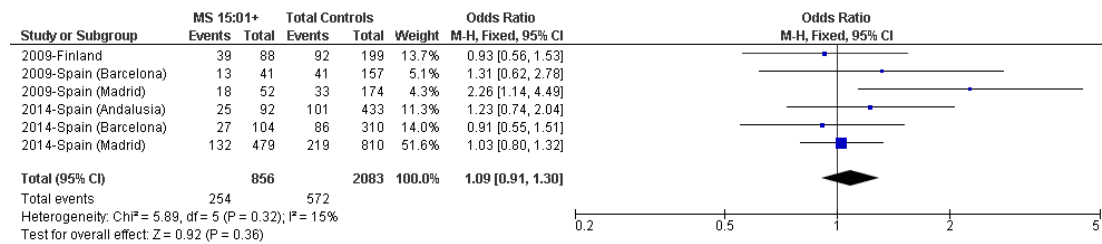


Fig. 3. Random effects meta-analysis including association studies of the *LILRA3* deletion with MS risk in *HLA-DRB1*15:01* carriers from populations of European descent. Study-specific ORs (squares) and 95% CIs (lines) were calculated for each individual dataset. Pooled OR and 95% CI were calculated combining all datasets. Del+/15:01+: carrier of *LILRA3* deletion and *HLA-DRB1*15:01* allele.

Analysis of the epistatic interaction between the *LILRA3* deletion and *HLA-DRB1*15:01* allele after gender-stratification

Further analysis was performed stratifying by gender the *LILRA3* deletion/ *HLA-DRB1*15:01* double-positive subjects and again, a significant difference was not detected

[Fig. 4, $p = 0.39$, OR (95% CI) = 1.09 (0.90-1.31)] with total homogeneity of strata ($I^2 = 0\%$). In addition, separated analyses in male and female groups were also performed, with similar results observed in both cases [male: $p = 0.76$, OR (95% CI) = 1.05 (0.76-1.47) and female: $p = 0.40$, OR (95% CI) = 1.10 (0.88-1.38)]. Neither significant evidence of sex-bias was observed in carriers of the *LILRA3* deletion lacking *HLA-DRB1*15:01* (data not shown).

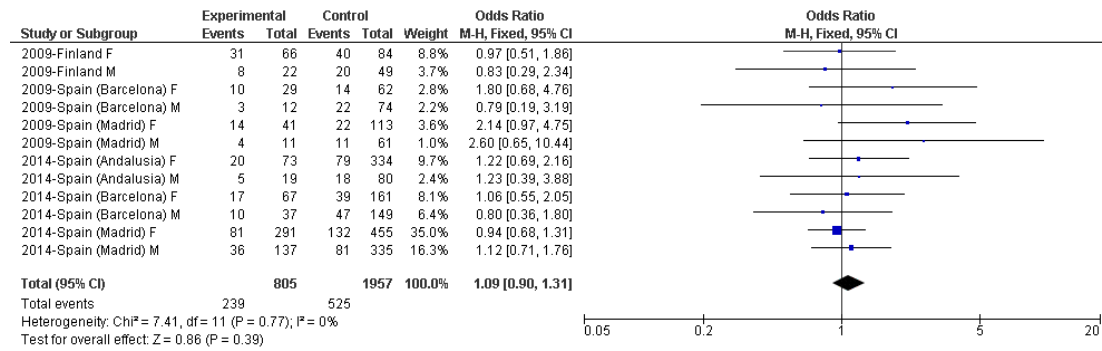


Fig. 4. Random effects meta-analysis of the association studies of the *LILRA3* deletion with MS risk in *HLA-DRB1*15:01* carriers from populations of European descent stratified by gender. Study-specific ORs (squares) and 95% CIs (lines) were calculated for each individual dataset. Pooled OR and 95% CI were calculated combining all datasets. Del+/15:01+: carrier of both *LILRA3* gene deletion and *HLA-DRB1*15:01* allele, M: male, F: female.

Discussion

The role of genetics in MS susceptibility is highly accepted and broadly described in the literature. At present, the greatest advances in the study of this genetic component have been achieved through GWAS, which analyze thousands of patients and ethnically matched controls searching for evidences to explain MS pathogenesis. However, this kind of studies does not explore genetic variants other than SNPs and the heritability underlying this and other complex diseases persists partially unexplained.

Previously reported studies tried to evaluate the influence of the *LILRA3* deletion on MS risk; however, their results were not fully concordant and it was difficult to draw a firm conclusion regarding the impact of this gene on MS. Two studies performed in German [21] and Spanish populations [20] showed a significant association of the *LILRA3* deletion with R-MS, with the strongest association found in the Spanish study. A MS cohort of French origin showed a high *LILRA3*-deletion frequency (46.1%), but the study included no matched French healthy controls[21], and it has not been included in our meta-analysis. Conversely, two studies performed in Finnish [6] and Polish populations [19] could not replicate the association. Therefore, a risk of publication bias does not seem to be present in the study of this gene. The meta-analysis combining the mentioned cohorts evidenced a statistically significant association of *LILRA3* deletion with MS risk [$p=0.03$, OR (95% CI)= 1.19 (1.02-1.40)], with modest evidence of heterogeneity ($I^2= 25\%$).

Considering that small sample sizes might explain the inconsistent results observed in the different reports, we aimed to ascertain the influence of the *LILRA3* deletion on R-MS susceptibility by studying this association in a large number of individuals from three different Spanish locations (Barcelona, Madrid and Andalusia) and performing a meta-analysis with the available data. Overall, our study with 6380 individuals (3257 MS

patients and 3123 healthy matched controls) did not reveal any significant difference when the distribution of the *LILRA3* deletion in MS patients and controls was compared [$p=0.25$, OR (95% CI)= 1.07 (0.95-1.19)], albeit in the context of a moderate evidence of heterogeneity ($I^2=32\%$). Substantial attenuation of the latter was achieved after the sensitivity analysis, which showed that the exclusion of the “2009-Spain (Madrid)” cohort reported by Ordonez *et al.* [20] led to a lower heterogeneity in the overall meta-analysis, still lacking significant association ($p=0.49$). Interestingly, the maximum reduction of heterogeneity was reached when the study with the highest impact on MS risk was removed from the meta-analysis, raising the possibility that the effect once observed was driven by a strong signal in this cohort. In conclusion, it seems that the *LILRA3* deletion by itself does not alter MS susceptibility.

Additionally, we also considered two other important factors: *HLA-DRB1*15:01* as the main risk allele described in MS and the putative gender bias, which might underlie the inconsistent results reported in the different studies. As previously described by Ordonez *et al.* [20], we assessed whether there was a synergistic effect between *HLA-DRB1*15:01* and the *LILRA3* deletion in order to confer MS risk, but no significant epistatic interaction was observed ($p=0.36$). Moreover, to evaluate whether the *LILRA3* deletion presented a possible gender bias in MS risk, as described by Du *et al.* in rheumatoid arthritis susceptibility [17], the epistatic interaction was also studied after gender stratification. Again, no significant differences were observed in male or female MS patients ($p=0.76$ and $p=0.40$, respectively), although two of the initial datasets were excluded as they lacked both gender- and *DRB1*15:01*- detailed information. Therefore, stratifications considering either the *HLA-DRB1*15:01* allele by itself or including also gender did not

evidence global association of *LILRA3* deletion with MS risk, demonstrating that a more stringent stratification of patients did not explain previous results.

Some works seemed to support the role of the *LILRA3* gene conferring risk to different autoimmune diseases and this gene would be another example of the well-established concept of shared genetic background among autoimmune diseases [14, 15, 17]. However, some of those studies warrant validation in independent cohorts or populations, provided that the studies including the greatest number of subjects have been performed in Asian populations. The frequency of the 6.7kb deletion differ greatly among ethnic groups, with extremely high values in Asians (0.56-0.84), and much lower values in Europeans (0.17) or Africans (0.10) [23]. It is well known that Asian and Caucasian populations show specific risk variants given the different genetic drift and the resulting different patterns of linkage disequilibrium.

Our results are concordant with those found in psoriasis, that reported lack of significant association between the *LILRA3* deletion and the disease in a European population [24], although this finding warrants replication as well. All these lines of evidence together showed that *LILRA3* does not seem to be underlying MS pathogenesis. Understanding the full load of genetic factors contributing to disease risk has potential for profound improvements in health care. Yet, the known genetic factors account for only a small portion of the estimated heritability of complex phenotypes and every effort to overcome this impasse should be emphasized.

Acknowledgements

We thank patients with multiple sclerosis and control subjects for making this study feasible. Financial support for the study was provided by: Fondo de Investigación Sanitaria (FIS)-Instituto de Salud Carlos III (ISCIII)-Fondos Europeos de Desarrollo Regional (FEDER) (grant numbers P12/00555; PI13/01527; PI13/02714; PI13/01466; PI13/0879 and RETICS-REEM RD12/0032/) and Junta de Andalucía (JA)- Fondos Europeos de Desarrollo Regional (FEDER) (grant number CTS2704). We thank Carlos Vilches (Immunogenetics & Histocompatibility, Instituto de Investigación Sanitaria Puerta de Hierro, Majadahonda, Madrid, Spain) for critical revision of the manuscript and enlightening comments.

Conflict of Interest Statement

The authors declare no conflict of interest.

References

1. Sawcer S, Franklin RJ, Ban M. Multiple sclerosis genetics. *Lancet Neurol*. 2014 May 19.
2. Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kempainen A, Cotsapas C, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet*. 2013 Nov;45(11):1353-60.
3. Anderson KJ, Allen RL. Regulation of T-cell immunity by leucocyte immunoglobulin-like receptors: innate immune receptors for self on antigen-presenting cells. *Immunology*. 2009 May;127(1):8-17.
4. Ryu M, Chen Y, Qi J, Liu J, Fan Z, Nam G, et al. LILRA3 binds both classical and non-classical HLA class I molecules but with reduced affinities compared to LILRB1/LILRB2: structural evidence. *PLoS One*. 2011;6(4):e19245.
5. Young NT, Canavez F, Uhrberg M, Shum BP, Parham P. Conserved organization of the ILT/LIR gene family within the polymorphic human leukocyte receptor complex. *Immunogenetics*. 2001 May-Jun;53(4):270-8.
6. Bonetti A, Koivisto K, Pirttila T, Elovaara I, Reunanen M, Laaksonen M, et al. A follow-up study of chromosome 19q13 in multiple sclerosis susceptibility. *J Neuroimmunol*. 2009 Mar 31;208(1-2):119-24.
7. Becker KG, Simon RM, Bailey-Wilson JE, Freidlin B, Biddison WE, McFarland HF, et al. Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci U S A*. 1998 Aug 18;95(17):9979-84.
8. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med*. 2007 Aug 30;357(9):851-62.
9. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. 2011 Aug 11;476(7359):214-9.
10. Golan D, Lander ES, Rosset S. Measuring missing heritability: Inferring the contribution of common variants. *Proc Natl Acad Sci U S A*. 2014 Nov 24.
11. Torkar M, Haude A, Milne S, Beck S, Trowsdale J, Wilson MJ. Arrangement of the ILT gene cluster: a common null allele of the ILT6 gene results from a 6.7-kbp deletion. *Eur J Immunol*. 2000 Dec;30(12):3655-62.
12. Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D, et al. Plasticity in the organization and sequences of human KIR/ILT gene families. *Proc Natl Acad Sci U S A*. 2000 Apr 25;97(9):4778-83.
13. Norman PJ, Carey BS, Stephens HA, Vaughan RW. DNA sequence variation and molecular genotyping of natural killer leukocyte immunoglobulin-like receptor, LILRA3. *Immunogenetics*. 2003 Jun;55(3):165-71.
14. Kabalak G, Dobberstein SB, Matthias T, Reuter S, The YH, Dorner T, et al. Association of immunoglobulin-like transcript 6 deficiency with Sjogren's syndrome. *Arthritis Rheum*. 2009 Oct;60(10):2923-5.

15. Du Y, Su Y, He J, Yang Y, Shi Y, Cui Y, et al. Impact of the leucocyte immunoglobulin-like receptor A3 (LILRA3) on susceptibility and subphenotypes of systemic lupus erythematosus and Sjogren's syndrome. *Ann Rheum Dis*. 2014 Jun 6.
16. An H, Chandra V, Piraino B, Borges L, Geczy C, McNeil HP, et al. Soluble LILRA3, a potential natural antiinflammatory protein, is increased in patients with rheumatoid arthritis and is tightly regulated by interleukin 10, tumor necrosis factor-alpha, and interferon-gamma. *J Rheumatol*. 2010 Aug 1;37(8):1596-606.
17. Du Y, Cui Y, Liu X, Hu F, Yang Y, Wu X, et al. Contribution of functional LILRA3, but not nonfunctional LILRA3, to sex bias in susceptibility and severity of anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheumatol*. 2014 Apr;66(4):822-30.
18. Thomas R, Matthias T, Witte T. Leukocyte immunoglobulin-like receptors as new players in autoimmunity. *Clin Rev Allergy Immunol*. 2010 Apr;38(2-3):159-62.
19. Wisniewski A, Wagner M, Nowak I, Bilinska M, Pokryszko-Dragan A, Jasek M, et al. 6.7-kbp deletion in LILRA3 (ILT6) gene is associated with later onset of the multiple sclerosis in a Polish population. *Hum Immunol*. 2013 Mar;74(3):353-7.
20. Ordonez D, Sanchez AJ, Martinez-Rodriguez JE, Cisneros E, Ramil E, Romo N, et al. Multiple sclerosis associates with LILRA3 deletion in Spanish patients. *Genes Immun*. 2009 Sep;10(6):579-85.
21. Koch S, Goedde R, Nigmatova V, Epplen JT, Muller N, de Seze J, et al. Association of multiple sclerosis with ILT6 deficiency. *Genes Immun*. 2005 Aug;6(5):445-7.
22. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*. 2001 Jul;50(1):121-7.
23. Hirayasu K, Ohashi J, Tanaka H, Kashiwase K, Ogawa A, Takanashi M, et al. Evidence for natural selection on leukocyte immunoglobulin-like receptors for HLA class I in Northeast Asians. *Am J Hum Genet*. 2008 May;82(5):1075-83.
24. Wisniewski A, Luszczyk W, Manczak M, Jasek M, Kubicka W, Cislo M, et al. Distribution of LILRA3 (ILT6/LIR4) deletion in psoriatic patients and healthy controls. *Hum Immunol*. 2003 Apr;64(4):458-61.